SRC TR 89-79

Distributed Detection From Multiple Sensors With Correlated Observations

By

Y.A. Chau and E. Geraniotis

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1989		**00-00-1989 to 00-00-1989**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Distributed Detection from Multiple Sensors with Correlated Observations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Maryland,Systems Research Center,College Park,MD,20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**46**	

DISTRIBUTED DETECTION FROM MULTIPLE SENSORS WITH CORRELATED OBSERVATIONS

Yawgeng A. Chau and Evaggelos Geraniotis

Department of Electrical Engineering
and Systems Research Center
University of Maryland
College Park, MD 20742

ABSTRACT

We address two problems of memoryless distributed detection with dependent observations across time and sensors. In the first problem, the observation sequence of each sensor consists of a common weak signal in additive dependent noise with stationary univariate and second-order joint densities; here the objective of the sensors is to cooperatively detect the presence of a weak signal. In the second problem, the observation sequence of each sensor is characterized by its stationary univariate and second-order joint densities; here the objective of the sensors is to cooperatively disrciminate between two arbitrary such sequences of observations. For both problems, the analysis and design are based on a common large sample size. The dependence across time and sensors is modeled by m-dependent, ϕ-mixing, or ρ-mixing processes. The performance of the two-sensor configuration for each problem is measured by an average cost, which couples the decisions of the sensors. The design criteria for the test satistics of the sensors, which constist of sums of memoryless nonlinearities, are established by using two-dimensional Chernoff bounds on the associated error probabilities involved in the average cost. The optimal nonlinearities are obtain as the solutions of linear coupled or uncoupled integral equations. Numerical results for specific cases of practical interest show that the performance of the proposed scheme is superior to the one which ignores the dependence across time and/or sensors for each of the two problems.

This research was supported in part by the Office of Naval Research under contract N00014-89-J-1375 and in part by the Systems Research Center at the University of Maryland, College Park, through the National Science Foundation's Engineering Research Centers Program: NSF CDR 8803012.

I. INTRODUCTION

Decentralized detection presents several interesting problems (see [1] and [2]). Classical theory of centralized hypothesis testing cannot be applied directly to practical systems of distributed decision makers (or sensors). The advantages of employing distributed sensors include low communication bandwidth, good reliability, and low cost.

In most of work on distributed decision making (see [1-5]), the observations (signal in noise) are assumed to be independent across time and/or sensors, an assumption that is intended to make the analysis tractable. However, the observations are generally dependent. Indeed, the observation sequences of the sensors become dependent across time individually for each sensor, when sampling rates increase, and correlated across sensors, when the locations of the sensors are close geographically.

In contrast to the model of independent observations, distributed detection from dependent observations across time and/or sensors are studied in this paper. In particular, two sensors without communication collect their own observations n times (the common sample size) and decide which hypothesis is true, H_0 or H_1. The object of each decision-making process is to minimize an average common cost, which consists of error probabilities and couples the two decisions in general. The form of the average cost is concerned with the consensus of the two sensors, and thus can be applied to the data fusion with a fixed rule such as 'OR', 'AND' or the majority rule.

Two problems of different descriptions for the observations are considered: in the first one, each observation consists of a weak signal in additive dependent noise with a stationary univariate density and second-order joint densities; in the second one, each observation itself is characterized by its univariate and second-order joint densities and thus, not necessary for the

signal in additive noise. For both types of observations, the analysis is based on a common large sample size for each sensor, which guarantees a high-quality performance.

The model of dependence used here is described by the mixing sequences, such as the m-dependent, the ϕ-mixing or the ρ-mixing sequences (see [6-8]). These models of dependence have been extensively and successfully used in single-sensor detection (see [9-11]). In general, the optimal detection scheme under the above model involves high-order (larger than two) densities, which are difficult to characterize, and needs huge space of memory for storing the dependent data. To avoid such inconvenience, we employ a suboptimal structure of memoryless nonlinearity for each sensor in this work (see [9]); each observation of sensor k ($k = 1, 2$) is passed through a nonlinearity $g_k(\cdot)$ and summed up to the sample size n to form the test statistic $T_{n,k}(g_k(\cdot))$. The activity of each sensor is to conduct a threshold test (as in the likelihood ratio test) by using its test statistic. The design criterion for each nonlinearity is maximizing an objective function of efficacy-type induced from the average cost by using the two-dimensional Chernoff bounds on error probabilities (see [9]). Each course of obtaining the optimal nonlinearity involves solving an integral equation.

Although the scheme of this paper is suboptimal, numerical results indicate that, with reasonable large sample size, the performance is better than that of the one which ignores the dependence.

The remainder of this paper is organized as follows: in Section II, we discuss the central limit theorem for the m-dependent, the ϕ-mixing and the ρ-mixing sequences and derive the two-dimensional Chernoff bounds. In Section III, the first description of observations for a weak signal in additive noise is considered. In Section IV, the second description of observations is considered. In Section V, numerical results from Cauchy noise for the first description of

observations (weak signal) and from Rayleigh vs. log normal (univariate and second-order joint) densities for the second description of observations (strong signal) are given. Conclusions are drawn in Section VI.

II. PRELIMINARIES

II.A. The Central Limit Theorem for Mixing Sequences

Let $\{X_i^{(k)}, i = 1, 2, \cdots, n\}$ be the observation sequence of sensor k, $(k = 1, 2)$; the test statistic of sensor k is given by

$$T_{n,k}(g_k) = \sum_{i=1}^{n} g_k(X_i^{(k)}) \tag{2}$$

where $g_k(\cdot)$ is the nonlinearity to be determined. The test of sensor k is

$$T_{n,k} \underset{H_0}{\overset{H_1}{\gtrless}} n\eta_k \tag{3}$$

where $n\eta_k$ is the threshold.

As mentioned in Section I, the model of dependence is one of the m-dependent, ϕ-mixing or ρ-mixing sequences. However, the model is an issue here, instead, it is essential that appropriate conditions are satisfied, as that, with a large sample size, the test statistic (3) of each sensor is asymptotically Gaussian distributed. To give an example of such "appropriate conditions," let us cite the central limit theorem for an arbitrary function h of the $\phi-$mixing sequences (see [6]).

Theorem 1: Suppose that $\{Y_j\}_{j=1}^{\infty}$ is a stationary $\phi-$mixing sequence with $\sum_{j=1}^{\infty} \phi_j^{\frac{1}{2}} < \infty$ and that h is a measurable function satisfying

$$E[h(Y_1)] = \mu, \quad var[h(Y_1)] < \infty$$

Then the series

$$\sigma^2(h) = var[h(Y_1)] + 2\sum_{j=1}^{m} cov[h(Y_1)h(Y_{j+1})] \tag{4}$$

converges absolutely as $m \to \infty$. Furthermore, if $\sigma^2 > 0$

$$\frac{1}{\sqrt{n}}[\sum_{j=1}^{n} h(Y_j) - n\mu]$$

converges in distribution to a Gaussian distribution with mean zero and variance σ^2.

Proof : Set $\tilde{h} = h - \mu_0$, thus $E[\tilde{h}(Y_1)] = 0$ and $E[\tilde{h}^2(Y_1)] < \infty$. The above Theorem follows according to the theorem on page 174 of [6].

In the following sections, we omit the arguments of the means $\mu_{k,i}(\cdot)$ and the variances $\sigma_{k,i}(\cdot)$, which are functionals of $g_k(\cdot)$ and thus functions of the observations. The m in the formulations is finite for m-dependent noise and $m \to \infty$, for ϕ-mixing or ρ- mixing noise.

II.B. Two-Dimensional Chernoff Bounds

The cost function used in this context has the form

$$C(d_1, d_2; u) = \begin{cases} 0 & \text{if } d_1 = d_2 = u \\ c_1 & \text{if } d_1 \neq d_2, \text{u=0} \\ c_1' & \text{if } d_1 \neq d_2, \text{u=1} \\ c_2 & \text{if } d_1 = d_2 \neq u, u = 0 \\ c_2' & \text{if } d_1 = d_2 \neq u, u = 1 \end{cases} \tag{5}$$

where d_k $(k = 1, 2)$ is the binary (0 or 1) decision of sensor k, u is the true phenomenon ($u = 0$ or 1), and c_1, c_1', c_2 and c_2' are non-negative constants. We also assume that $c_2 \geq c_1, c_1'$ and $c_2' \geq c_1, c_1'$, i.e., the penalty is higher when both sensors make wrong decisions. This form of cost is of the Bayesian type and has been useful in [1]. The average cost, which couples the decisions of the two sensors is

$$\begin{aligned} & E[C(d_1, d_2; u)] = pE_0[C(d_1, d_2; u)] + (1-p)E_1[C(d_1, d_2; u)] \\ = \; & p\{c_1[P_0(1,0) + P_0(0,1)] + c_2 P_0(1,1)\} + (1-p)\{c_1'[P_1(1,0) + P_1(0,1)] + c_2' P_1(0,0)\} \end{aligned} \tag{6}$$

where p is the a prior probability of H_0, P_i $(i = 0, 1)$ are error probabilities under H_i, $P_i(1,0) = P_i(T_{n,1} > n\eta_1, T_{n,2} \leq n\eta_2)$, $P_i(0,1) = P_i(T_{n,1} \leq n\eta_1, T_{n,2} > n\eta_2)$ $(i = 0, 1)$, $P_0(1,1) = P_0(T_{n,1} > n\eta_1, T_{n,2} > n\eta_2)$ and $P_1(0,0) = P_1(T_{n,1} \leq n\eta_1, T_{n,2} \leq n\eta_2)$.

Because of the complexity of the average cost, there is no directly tractable approach by which a nice criterion can be reached for each optimal nonlinearity g_k $(k = 1, 2)$. Instead of optimizing (6) directly, we first derive the two-dimensional Chernoff bounds on the error probabilities in (6). In Sections III and IV, we approximate the bounds under a large sample size. Our purpose is to approximate and combine the terms in (6) as an exponential form, i.e., $E[C(d_1(g_1), d_2(g_2); u)] \approx A \cdot \exp[J(g_1, g_2)]$, so that we have a closed and meaningful form of design criterion for each nonlinearity.

Define

$$Y(\eta_1, \eta_2) = \begin{cases} 1 & \text{if } T_{n,1} > n\eta_1, T_{n,2} > n\eta_2 \\ 0 & \text{otherwise,} \end{cases}$$

where we assume that, for $k = 1, 2$, $T_{n,k} \sim \mathcal{N}(n\mu_{k,i}, n\sigma^2_{k,i})$ under H_i, $(i = 0, 1)$, and the pair $(T_{n,1}, T_{n,2})$ is bivariate Gaussian distributed with the correlation coefficient $\rho_i = E_i[T_{n,1} \cdot T_{n,2}]$. Then for $s_1 \geq 0, s_2 \geq 0$

$$e^{s_1 n\eta_1 + s_2 n\eta_2} Y(\eta_1, \eta_2) \leq e^{s_1 T_1 + s_2 T_2}. \tag{7}$$

Taking the expectation of (7) under H_0 and we have, after some manipulations

$$\begin{aligned} &P_0(T_{n,1} > n\eta_1, T_{n,2} > n\eta_2) \leq e^{-s_1 n\eta_1 - s_2 n\eta_2} E_0[e^{s_1 T_{n,1} + s_2 T_{n,2}}] \\ &= \sqrt{2\pi} \exp\left[\frac{n\sigma^2_{1,0} s_1^2}{2} + \frac{n\sigma^2_{2,0} s_2^2}{2} + n\rho_0 \sigma_{1,0}\sigma_{2,0} s_1 s_2 - s_1 n(\eta_1 - \mu_{1,0}) - s_2 n(\eta_2 - \mu_{2,0})\right]. \end{aligned} \tag{8}$$

Similarly, for $t_1 \geq 0, t_2 \geq 0$, and $q_{1,i} \geq 0, q_{2,i} \geq 0$ and $r_{1,i} \geq 0, r_{2,i} \geq 0$ we have

$$\begin{aligned} &P_1(T_{n,1} \leq n\eta_1, T_{n,2} \leq n\eta_2) \leq e^{t_1 n\eta_1 + t_2 n\eta_2} E_1[e^{-t_1 T_{n,1} - t_2 T_{n,2}}] \\ &= \sqrt{2\pi} \exp\left[\frac{n\sigma^2_{1,1} s_1^2}{2} + \frac{n\sigma^2_{2,1} s_2^2}{2} + n\rho_1 \sigma_{1,1}\sigma_{2,1} s_1 s_2 + s_1 n(\eta_1 - \mu_{1,1}) + s_2 n(\eta_2 - \mu_{2,1})\right] \end{aligned} \tag{9}$$

under H_1,

$$P_i(T_{n,1} > n\eta_1, T_{n,2} \leq n\eta_2) \leq e^{-q_{1,i} n\eta_1 + q_{2,i} n\eta_2} E_i[e^{q_{1,i} T_{n,1} - q_{2,i} T_{n,2}}]$$

$$= \sqrt{2\pi}\exp\left[\frac{n\sigma_{1,i}1^2q_{1,i}^2}{2}+\frac{n\sigma_{2,i}^2q_{2,i}^2}{2}-n\rho_i\sigma_{1,i}\sigma_{2,i}q_{1,i}q_{2,i}-q_{1,i}n(\eta_1-\mu_{1,i})+q_{2,i}n(\eta_2-\mu_{2,i})\right] \tag{10}$$

under H_i, $(i=0,1)$ and

$$P_i(T_{n,1}\le n\eta_1, T_{n,2}>n\eta_2) \le e^{r_{1,i}n\eta_1-r_{2,i}n\eta_2}E_i[e^{-r_{1,i}T_{n,1}+r_{2,i}T_{n,2}}]$$
$$= \sqrt{2\pi}\exp\left[\frac{n\sigma_{1,i}^2r_{1,i}^2}{2}+\frac{n\sigma_{2,i}^2r_{2,i}^2}{2}-n\rho_i\sigma_{1,i}\sigma_{2,i}r_{1,i}r_{2,i}+r_{1,i}n(\eta_1-\mu_{1,i})-r_{2,i}n(\eta_2-\mu_{2,i})\right] \tag{11}$$

under $H_i, i=0,1$. The bounds given by (8)-(11) supply good approximations to the error probabilities in the average cost, as the sample size increases. The positive constants $s_k, t_k, q_{k,i}, r_{k,i}$ are to be adjusted to set final forms of the bounds. Since the weights (c_2, c_2') of $P_0(1,1)$ or $P_1(0,0)$ are greater than that (c_1, c_1') of $P_i(0,1)$ or $P_i(1,0)$ $(i=0,1)$, we first minimize the bounds in (8) and (9) with respect to (s_1,s_2) and (t_1,t_2), respectively. If we minimize (8) with respect to (s_1,s_2), we have

$$P_0(T_{n,1}>n\eta_1, T_{n,2}>n\eta_2)$$
$$\le \sqrt{2\pi}\exp\left\{-\frac{n}{2(1-\rho_0^2)}\left[\frac{(\eta_1-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,0})^2}{\sigma_{2,0}^2}-\frac{2\rho_0(\eta_1-\mu_{1,0})(\eta_2-\mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right]\right\} \tag{12}$$

and the corresponding (s_1,s_2) are

$$\tilde{s}_1=\frac{(\eta_1-\mu_{1,0})\sigma_{2,0}-(\eta_2-\mu_{2,0})\rho_0\sigma_1}{(1-\rho_0^2)\sigma_{1,0}^2\sigma_{2,0}} \tag{13}$$

$$\tilde{s}_2=\frac{(\eta_2-\mu_{2,0})\sigma_{1,0}-(\eta_1-\mu_{1,0})\rho_1\sigma_{2,0}}{(1-\rho_0^2)\sigma_{2,0}^2\sigma_{1,0}}, \tag{14}$$

where $0\le\rho_0^2\le 1$. Similarly, if we minimize (13) with respect to (t_1,t_2), we have

$$P_1(T_{n,1}\le n\eta_1, T_{n,2}\le n\eta_2)$$
$$\le \sqrt{2\pi}\exp\left\{-\frac{n}{2(1-\rho_1^2)}\left[\frac{(\mu_{1,1}-\eta_1)^2}{\sigma_{1,1}^2}+\frac{(\mu_{2,1}-\eta_2)^2}{\sigma_{2,1}^2}-\frac{2\rho_1(\mu_{1,1}-\eta_1)(\mu_{2,1}-\eta_2)}{\sigma_{1,1}\sigma_{2,1}}\right]\right\} \tag{15}$$

and the corresponding (t_1, t_2) are

$$\tilde{t}_1 = \frac{(\mu_{1,1} - \eta_1)\sigma_{2,1} - (\mu_{2,1} - \eta_2)\rho_1\sigma_{1,1}}{(1-\rho_1^2){\sigma_{1,1}}^2\sigma_{2,1}} \tag{16}$$

$$\tilde{t}_2 = \frac{(\mu_{2,1} - \eta_2)\sigma_{1,1} - (\mu_{1,1} - \eta_1)\rho_1\sigma_{2,1}}{(1-\rho_1^2){\sigma_{2,1}}^2\sigma_{1,1}} \tag{17}$$

Under the restrictions $s_k \geq 0$ and $t_k \geq 0$ $(k = 1, 2)$, we have

$$\frac{\eta_1 - \mu_{1,0}}{\sigma_{1,0}} > \frac{\eta_2 - \mu_{2,0}}{\sigma_{2,0}}\rho_0, \tag{18}$$

$$\frac{\eta_2 - \mu_{2,0}}{\sigma_{2,0}} > \frac{\eta_1 - \mu_{1,0}}{\sigma_{1,0}}\rho_0, \tag{19}$$

$$\frac{\mu_{1,1} - \eta_1}{\sigma_{1,1}} > \frac{\mu_{2,1} - \eta_2}{\sigma_{2,1}}\rho_1 \tag{20}$$

and

$$\frac{\mu_{2,1} - \eta_2}{\sigma_{2,1}} > \frac{\mu_{1,1} - \eta_1}{\sigma_{1,1}}\rho_1. \tag{21}$$

Let us now consider (10) and (11). We can minimize (10) with respect to $(r_{1,i}, r_{2,i})$, or (11) with respect to $(q_{1,i}, q_{2,i})$. However, either minimization will lead to results contradictory to (18)-(19) or (20)-(21), if the constraint $r_{k,i} \geq 0$ or $q_{k,i} \geq 0$ $(i = 0, 1; k = 1, 2)$ are to be satisfied. Therefore, we choose

$$r_{0,1} = \frac{\eta_1 - \mu_{1,0}}{\sigma_{1,0}^2}, \quad r_{2,0} = 0 \tag{22}$$

$$r_{2,1} = \frac{\mu_{2,1} - \eta_1}{\sigma_{2,1}^2}, \quad r_{1,1} = 0 \tag{23}$$

$$q_{2,0} = \frac{\eta_1 - \mu_{2,0}}{\sigma_{2,0}^2}, \quad q_{1,0} = 0 \tag{24}$$

and

$$q_{1,1} = \frac{\mu_{1,1} - \eta_1}{\sigma_{1,1}^2}, \quad q_{2,1} = 0 \tag{25}$$

which give the bounds on (10) and (11) as

$$P_i(T_{n,1} > n\eta_1, T_{n,2} \leq n\eta_2) \leq \sqrt{2\pi}\exp\left[\frac{-n(\eta_1 - \mu_{1,i})^2}{2\sigma_{1,i}^2}\right] \tag{26}$$

and

$$P_i(T_{n,1} \leq n\eta_1, T_{n,2} > n\eta_2) \leq \sqrt{2\pi} \exp\left[\frac{-n(\eta_2 - \mu_{2,i})^2}{2\sigma_{2,i}^2}\right]. \tag{27}$$

In the above bounds, all parameters $s_k, t_k, r_{k,i}$ and $q_{k,i}$ ($i = 0, 1$ and $k = 1, 2$) are nonnegative. The constraints satisfied by $\eta_k, \mu_{k,i}$ and $\sigma_{k,i}$ are described by (18)-(21) and

$$\mu_{k,0} \leq \eta_k \leq \mu_{k,1} \tag{28}$$

for $k = 1, 2$.

Finally the bound on the average cost has the form

$$\begin{aligned}
&E[C(d_1, d_2; u)] \leq B_u \\
&= p\{c_1\left\{\sqrt{2\pi}\exp\left[\frac{-n(\eta_1 - \mu_{1,0})^2}{2\sigma_{1,0}^2}\right] + \sqrt{2\pi}\exp\left[\frac{-n(\eta_2 - \mu_{2,0})^2}{2\sigma_{2,0}^2}\right]\right\} \\
&\quad + c_2\sqrt{2\pi}\exp\left\{-\frac{n}{2(1-\rho_0^2)}\left[\frac{(\eta_1 - \mu_{1,0})^2}{\sigma_{1,0}^2} + \frac{(\eta_2 - \mu_{2,0})^2}{{\sigma_{2,0}}^2} - \frac{2\rho_0(\eta_1 - \mu_{1,0})(\eta_2 - \mu_{2,0})}{\sigma_{1,0}\sigma_{2,0}}\right]\right\}\} \\
&+ (1-p)\{c_1'\left\{\sqrt{2\pi}\exp\left[\frac{-n(\eta_1 - \mu_{1,1})^2}{2\sigma_{1,1}^2}\right] + \sqrt{2\pi}\exp\left[\frac{-n(\eta_2 - \mu_{2,1})^2}{2\sigma_{2,1}^2}\right]\right\} \\
&\quad + c_2'\sqrt{2\pi}\exp\left\{-\frac{n}{2(1-\rho_1^2)}\left[\frac{(\eta_1 - \mu_{1,1})^2}{\sigma_{1,1}^2} + \frac{(\eta_2 - \mu_{2,1})^2}{\sigma_{2,1}^2} - \frac{2\rho_1(\eta_1 - \mu_{1,1})(\eta_2 - \mu_{2,1})}{\sigma_{1,1}\sigma_{2,1}}\right]\right\}\}
\end{aligned} \tag{29}$$

For the purpose of convenience, in the following sections we will set $c_1 = c_1'$ and $c_2 = c_2'$, which can be applied to the data fusion with a majority fusion-rule. For other fusion-rules, the techniques in the following sections remains useful with minor modifications.

III. WEAK SIGNALS

We consider the following observation model of a weak signal in two-sensor detection

$$\begin{aligned} H_0^{(k)} &: X_i^{(k)} = N_i^{(k)} \\ H_1^{(k)} &: X_i^{(k)} = \theta + N_i^{(k)}, \ i = 1, \cdots, n; \ k = 1, 2 \end{aligned} \tag{30}$$

where $\theta \to 0$ with $\theta > 0$ is the nonrandom weak signal and $\{N_i^{(k)}, i = 1, \cdots, n\}$, $(k = 1, 2)$ are two noise sequences with stationary univariate and second-order joint densities. The asymptotic analysis based on a large sample size, which guarantees a high-quality performance for $\theta \to \infty$, is used. For this model of observations, we assume that

$$E_\theta[g_k] < \infty, \ \ var_\theta(g_k) < \infty \tag{31}$$

and

$$\sigma^2_{k,\theta}(g_k) = var_{k,\theta}[g_k(X_i^{(k)})] + 2\sum_{j=1}^{m} cov_{k,\theta}[g_k(X_i^{(k)})g_k(X_{j+1}^{(k)})] > 0 \tag{32}$$

for all $\theta \geq 0$. In addition, in this description of observations we use the notations

$$\sigma^2_{k,\theta}(g_k) = \sigma^2_{k,1}(g_k), \tag{33}$$

$$n\rho_\theta(g_1, g_2) = n\rho_1(g_1, g_2) = E_\theta[(T_{n,1} - n\mu_{1,i})(T_{n,2} - n\mu_{2,i})] \tag{34}$$

and

$$\mu_{k,\theta}(g_k) = \mu_{k,1}(g_k) = \int g_k(x) f_k(x - \theta) dx \tag{35}$$

for $\theta > 0$, where $f_k(\cdot)$ the univariate density of $N_i^{(k)}$ and thus of $X_i^{(k)}$ under H_0.

As discussed in Section II, conditioned on H_i for $I = 0, 1$ and $\theta \geq 0$, the test statistics $T_{n,k}$ of each sensor are asymptotically Gaussian distributed with means $n\mu_{k,\theta}(g_k)$ and variances $n\sigma^2_{k,\theta}(g_k)$. Furthermore, $T_{n,1}$ and $T_{n,2}$ are bivariate Gaussian distributed with the correlation

coefficient $n\rho_\theta(g_1, g_2)$. In addition, the following regularity conditions are assumed, for $k = 1, 2$ and $\theta \to 0$, under a large sample size

$$\frac{\partial}{\partial\theta}[\int g_k(x) f_k(x-\theta)dx] = \int \frac{\partial}{\partial\theta}[g_k(x) f_k(x-\theta)]dx, \tag{36}$$

$$\frac{\mu_{k,\theta}(g_k) - \mu_{k,0}(g_k)}{\theta} \to \frac{\partial\mu_\theta(g_k)}{\partial\theta}|_{\theta=0} > 0, \tag{37}$$

$$\sigma_{k,\theta}(g_k) \to \sigma_{k,0}(g_k), \tag{38}$$

and

$$\rho_\theta(g_1, g_2) \to \rho_0(g_1, g_2). \tag{39}$$

Since $\theta \to \infty$, we use $\sigma_{k,0}$ and ρ_0 to represent the variance and the correlation coefficient under either hypothesis in the following Sections.

III.A. Dependence Across Time

First let us consider the case when the two observation sequences are independent. The locally optimal test of each sensor as shown in [1] is a likelihood ratio test with a data-dependent threshold. However, the computation of the optimal thresholds is generally complicated and depends on high-order (larger than two) densities of the observation sequences that are difficult to characterized. Here we develop a suboptimal approach by using the central limit theorem and the bounds derived in the Section II. The threshold of each sensor for design purpose is easy to determine and depends only on the means of the nonlinearity.

Since the correlation coefficient ρ_0 in Section II is zero in this case, the bound of (29) in Section II has the form

$$\begin{aligned} B_u \;=\; & p\{c_1\sqrt{2\pi}\exp\left[\frac{-n(\eta_1-\mu_{1,0})^2}{2\sigma_{1,0}^2}\right] + c_1\sqrt{2\pi}\exp\left[\frac{-n(\eta_2-\mu_{2,0})^2}{2\sigma_{2,0}^2}\right] \\ & + c_2\sqrt{2\pi}\exp\left[\frac{-n(\eta_1-\mu_{1,0})^2}{2\sigma_{1,0}^2} + \frac{-n(\eta_2-\mu_{2,0})^2}{2\sigma_{2,0}^2}\right]\} \end{aligned}$$

$$+(1-p)\{c_1\sqrt{2\pi}\exp\left[\frac{-n(\eta_1-\mu_{1,1})^2}{2\sigma_{1,0}^2}\right]+c_1\sqrt{2\pi}\exp\left[\frac{-n(\eta_2-\mu_{2,1})^2}{2\sigma_{2,0}^2}\right]$$
$$+c_2\sqrt{2\pi}\exp\left[\frac{-n(\eta_1-\mu_{1,1})^2}{2\sigma_{1,0}^2}+\frac{-n(\eta_2-\mu_{2,1})^2}{2\sigma_{1,0}^2}\right]\} \tag{40}$$

where $\mu_{k,0} \leq \eta_k \leq \mu_{k,1}$. Notice that there are two parts in (40): one is from the error probability under H_0 and the other is from the error probability under H_1. Using the fact that for positive constants (a_1, b_1) and negative constants (a_2, b_2)

$$a_1 e^{na_2} + b_1 e^{nb_2} \approx 2a_1 b_1 e^{(na_2+nb_2)/2} \tag{41}$$

as n increases, we can approximate the bound in (40) by

$$B_u \approx pA\exp\left\{\frac{-3n}{8}\left[\frac{(\eta_1-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,0})^2}{\sigma_{2,0}^2}\right]\right\}$$
$$+(1-p)A\exp\left\{\frac{-3n}{8}\left[\frac{(\eta_1-\mu_{1,\theta})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,\theta})^2}{\sigma_{2,0}^2}\right]\right\} \tag{42}$$

where A is a known constant. Because the optimal thresholds are difficult to calculate, we use the suboptimal thresholds determined in an *ad hoc* way. Since the error probabilities cannot be known *a priori*, we pose the condition

$$\frac{(\eta_1-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,0})^2}{\sigma_{2,0}^2}=\frac{(\eta_1-\mu_{1,\theta})^2}{\sigma_{1,0}^2}+\frac{(\eta_2-\mu_{2,\theta})^2}{\sigma_{2,0}^2} \tag{43}$$

for all $\mu_{k,0}$, $\mu_{k,\theta}$ and $\sigma_{k,0}$ $(k = 1, 2)$ to balance the two error probabilities under two hypotheses in (40), so that we do not bias either hypothesis, One solution of (43) for (η_1, η_2) is

$$\eta_k = \frac{\mu_{k,0}+\mu_{k,\theta}}{2}, \quad k = 1, 2. \tag{44}$$

It is easy to check that the constraints (18)-(21) and (28) are satisfied by using the above thresholds. The bound (42) employing η_k $(k = 1, 2)$ given by (44) is

$$B_u \approx A\exp\left\{\frac{-3n}{32}\left[\frac{(\mu_{1,\theta}-\mu_{1,0})^2}{\sigma_{1,0}^2}+\frac{(\mu_{2,\theta}-\mu_{2,0})^2}{\sigma_{2,0}^2}\right]\right\} \tag{45}$$

We have reached the exponential form given by (45) for the approximate upper bound on the average cost by using the above thresholds. Minimizing this exponential form with respect to (g_1, g_2) is equivalent to maximizing the exponent

$$\begin{aligned} & \frac{n(\mu_{1,\theta}(g_1) - \mu_{1,0}(g_1))^2}{\sigma^2_{1,0}(g_1)} + \frac{n(\mu_{2,\theta}(g_2) - \mu_{2,0}(g_2))^2}{\sigma^2_{2,0}(g_2)} \\ = \quad & \theta^2 \left\{ \frac{[(\mu_{1,\theta}(g_1) - \mu_{1,0}(g_1))/\theta]^2}{\sigma^2_{1,0}(g_1)} + \frac{[(\mu_{2,\theta}(g_2) - \mu_{2,0}(g_2))/\theta]^2}{\sigma^2_{2,0}(g_2)} \right\} \end{aligned} \tag{46}$$

where the weak signal $\theta \to 0$. Using the regularity condition (36) and normalizing (46) by dividing it by θ^2, we have the equivalent objective function for design criterion

$$J(g_1, g_2) = \frac{\{\partial E_\theta[g_1]/\partial\theta\}^2_{\theta=0}}{\sigma^2_{1,0}(g_1)} + \frac{\{\partial E_\theta[g_2]/\partial\theta\}^2_{\theta=0}}{\sigma^2_{2,0}(g_2)} \tag{47}$$

which is to be maximized with respect to (g_1, g_2). We notice that (47) is the summation of two efficacies, and thus the criterion becomes a generalization of the ARE in single-sensor detection (see [9]). In fact, we have shown the connection between the general ARE and the average cost for the detection of a weak signal in a multiple-sensor environment. Because the observation sequences of the two sensors are independent of each other, we can maximize the efficacy for each sensor in (47) with respect to its g_k. Notice that the efficacy for each sensor in (47) is invariant under the scaling of its g_k. Using the same technique as in [9] we have the following integral equation

$$-f'_k(x)/f_k(x) - \int_{-\infty}^{\infty} K_k(x, y) g_k(y) dy = g_k(x) \tag{48}$$

for optimal g_k, where the integration kernel $K_k(x, y)$ has the form

$$K_k(x, y) = 2 \sum_{j=1}^{m} f^{(k)}_{N_{j+1}/N_1}(y|x) - (2m+1) f_k(y), \; k = 1, 2. \tag{49}$$

The procedure of solving (48) was addressed in [9] and will not be duplicated here.

III.B. Dependence Across Time and Sensors

We consider now the case when the two stationary observation sequences are dependent across time and sensors. The situation of interest here is that the univariate and second-order joint densities of the two observation sequences are identical and the structures (i.e., the non-linearities and thresholds) of the two sensors are the same. In our formulations, we use the notations $f_1(\cdot) = f_2(\cdot) = f(\cdot)$, $f_j^{(1)}(\cdot,\cdot) = f_j^{(2)}(\cdot,\cdot) = f_j(\cdot,\cdot)$, $g_1(\cdot) = g_2(\cdot) = g(\cdot)$ and $\eta_1 = \eta_2 = \eta$. Consequently, $\mu_{1,0} = \mu_{2,0} = \mu_0$, $\mu_{1,\theta} = \mu_{2,\theta} = \mu_\theta$, and $\sigma_{1,0} = \sigma_{2,0} = \sigma_0$. These conditions characterize a model, for which the channels of the two sensors are symmetric and the two sensors are implemented under the consideration of regulation with the same structure.

Under the above symmetric conditions, the bound described by (29) has the form

$$\begin{aligned} B_u &= p\{2\sqrt{2\pi}c_1 \exp[\frac{-n(\eta-\mu_0)^2}{2\sigma_0^2}] + \sqrt{2\pi}c_2 \exp[\frac{-n(\eta-\mu_0)^2}{(1+\rho_0)\sigma_0^2}]\} \\ &\quad +(1-p)\{2\sqrt{2\pi}c_1 \exp[\frac{-n(\eta-\mu_\theta)^2}{2\sigma_0^2}] + \sqrt{2\pi}c_2 \exp[\frac{-n(\eta-\mu_\theta)^2}{(1+\rho_0)\sigma_0^2}]\} \end{aligned} \tag{50}$$

Using the same technique of balance condition as in the previous case

$$\frac{-n(\eta-\mu_0)^2}{2\sigma_0^2} = \frac{-n(\eta-\mu_\theta)^2}{2\sigma_0^2} \tag{51}$$

$$\frac{-n(\eta-\mu_0)^2}{(1+\rho_0)\sigma_0^2} = \frac{-n(\eta-\mu_\theta)^2}{(1+\rho_0)\sigma_0^2}, \tag{52}$$

we obtain one solution for η

$$\eta = \frac{\mu_0+\mu_\theta}{2}. \tag{53}$$

Again conditions (18)-(21) and (28) are satisfied with this threshold. The bound defined from this threshold is

$$B_u = 2\sqrt{2\pi}c_1 \exp\left\{\frac{-n(\mu_\theta-\mu_0)^2}{8\sigma_0^2}\right\} + \sqrt{2\pi}c_2 \exp\left\{\frac{-n(\mu_\theta-\mu_0)^2}{4(1+\rho_0)\sigma_0^2}\right\}. \tag{54}$$

The globally optimal nonlinearity that minimizes (54) is difficult to obtain and may even not exist. Therefore, we design the nonlinearity in the *ad hoc* way described as follows.

We notice that there are two exponents in (54), namely

$$R_1(g) = \frac{n(\mu_\theta(g) - \mu_0(g))^2}{2\sigma_0^2(g)} = n\theta^2 \frac{[(\mu_\theta(g) - \mu_0(g))/\theta]^2}{8\sigma_0^2(g)} = n\theta^2 \frac{\{[\partial E_\theta[g]/\partial\theta]_{\theta=0}\}^2}{8\sigma_0^2(g)} \tag{55}$$

and

$$R_2(g) = \frac{n(\mu_\theta(g) - \mu_0(g))^2}{(1+\rho_0(g))\sigma_0^2(g)} = n\theta^2 \frac{[(\mu_\theta(g) - \mu_0(g))/\theta]^2}{4(1+\rho_0(g))\sigma_0^2(g)} = n\theta^2 \frac{\{[\partial E[g]/\partial\theta]_{\theta=0}\}^2}{4(1+\rho_0(g))\sigma_0^2(g)} \tag{56}$$

as $\theta \to 0$. First we obtain the optimal $g_{R_1}(\cdot)$ by maximizing the functional $R_1(\cdot)$ with respect to $g(\cdot)$. Then we maximize $R_2(\cdot)$ to obtain the optimal $g_{R_2}(\cdot)$. From $g_{R_1}(\cdot)$ and $g_{R_2}(\cdot)$ we compose the nonlinearity of each sensor as

$$g(x) = \tau_w g_{R_1}(x) + (1 - \tau_w) g_{R_2}(x), \tag{57}$$

where the parameter τ_w is determined through numerical experimentation. The optimization of $R_1(\cdot)$ has been discussed in Subsection III.A. For the optimization of $R_2(\cdot)$, we can derive an integral equation satisfied by the optimal nonlinearity $g_{R_2}(\cdot)$ in a manner similar to that in [9].

Notice that, for all $\alpha \neq 0$, $R_2(g) = R_2(\alpha g)$, i.e., $R_2(\cdot)$ is invariant under the scaling of its $g(\cdot)$. Therefore under the regularity condition (37), maximizing $R_2(\cdot)$ with respect to $g(\cdot)$ is equivalent to maximizing $-\int g(x)f'(x)dx$ under the constraint that $(1+\rho_0(g))\sigma_0^2(g)$is a constant. Thus the optimization of $R_2(g)$ is characterized by $\max_g H(g)$ with

$$H(g) = \{-\int g(x) f_0'(x) dx + \lambda[1+\rho_0(g)]\sigma_0^2(g)\} \tag{58}$$

where λ is the Lagrange multiplier and

$$\begin{aligned} \rho_0(g)\sigma_0^2(g) &= E_0\left\{\frac{[T_{n,1} - n\mu_0]}{\sqrt{n}} \cdot \frac{[T_{n,2} - n\mu_0]}{\sqrt{n}}\right\} \\ &= E[g(x_1^{(1)})g(x_1^{(2)})] + 2\sum_{j=1}^{m} E[g(x_1^{(1)})g(x_{j+1}^{(2)})] - (2m+1)\mu_0^2 \end{aligned} \tag{59}$$

by using a large sample size.

Denote the variation of $g(\cdot)$ by $\delta g(\cdot)$ and let $J_g(\epsilon) = H(g + \epsilon\delta g)$. Then the necessary and sufficient conditions satisfied by the optimal g for arbitrary δg is

$$J'_g(0) = 0 \tag{60}$$

and

$$J''_g(0) < 0. \tag{61}$$

Let $f_{N_1^{(1)},N_{j+1}^{(1)}}$ be the joint density of $N_1^{(1)}$ and $N_{j+1}^{(1)}$ and $\tilde{f}_{N_1^{(1)},N_{j+1}^{(2)}}$ the joint density of $N_1^{(1)}$ and $N_{j+1}^{(2)}$. We have

$$J'_g(0)=\int \{-f'(x) + 2\lambda[g(x)f(x) + \int [\sum_{j=1}^{m}(f_j(x,y) + \tilde{f}_j(x,y)) - 2f(x)f(y)]dy]\}\delta g(x)dx \tag{62}$$

where

$$\begin{aligned}\tilde{f}_j(x,y) &= (\tilde{f}_{N_1^{(1)},N_2^{(2)}}(x,y) + \tilde{f}_{N_1^{(1)},N_2^{(2)}}(y,x))/2 \\ &\quad + \tilde{f}_{N_1^{(1)},N_{j+1}^{(2)}}(x,y) + \tilde{f}_{N_1^{(1)},N_{j+1}^{(2)}}(y,x) - 2f(x)f(y)\end{aligned} \tag{63}$$

and

$$f_j(x,y) = f_{N_1^{(1)},N_{j+1}^{(1)}}(x,y) + f_{N_1^{(1)},N_{j+1}^{(1)}}(y,x) - 2f(x)f(y). \tag{64}$$

From (60) and (61), we have the linear integral equation

$$-f'(x) + 2\lambda[g(x)f(x) + \int_{-\infty}^{\infty} g(y)(\sum_{j=1}^{m} f_j(x,y) + \tilde{f}_j(x,y))dy] = 0 \tag{65}$$

where λ is a scaling factor of g. From (61) we obtain the sufficient condition

$$2\lambda(1 + \rho_0(\delta g))\sigma_0^2(\delta g) < 0. \tag{66}$$

Since $\rho_0(.) < 1$, (66) holds for negative λ. Following the tradition in single-sensor detection (see [9]) for i.i.d. observations, we set $\lambda = -1/2$; thus the integral equation to be solved for the

optimal g can be expressed as

$$-f'(x)/f(x) - \int_{-\infty}^{\infty} K_c(x,y)g(y)dy = g(x), \tag{67}$$

where $g_{lo}(x) = -f'(x)/f(x)$ is the locally-optimum detector for i.i.d. observations across time and sensors and

$$\begin{aligned} K_c(x,y) &= (\tilde{f}_{N_1^{(2)}/N_1^{(1)}}(y|x) + \tilde{f}_{N_1^{(1)}/N_1^{(2)}}(y|x))/2 - 2f(y) \\ &\quad + \sum_{j=1}^{m}[f_{N_{j+1}^{(1)}/N_1^{(1)}}(y|x) + f_{N_1^{(1)}/N_{j+1}^{(1)}}(y|x) \\ &\quad\quad + \tilde{f}_{N_{j+1}^{(2)}/N_1^{(1)}}(y|x) + \tilde{f}_{N_{j+1}^{(1)}/N_1^{(2)}}(y|x) - 4f(y)] \\ &= \tilde{f}_{N_1^{(2)}/N_1^{(1)}}(y|x) - 2f(y) + 2\sum_{j=1}^{m}[f_{N_{j+1}^{(1)}/N_1^{(1)}}(y|x) + \tilde{f}_{N_{j+1}^{(2)}/N_1^{(1)}}(y|x) - 2f(y)] \end{aligned} \tag{68}$$

is the kernel of integration.

IV. STRONG SIGNALS

In this section we consider the second description of observations, which is characterized by the univariate and second-order joint densities. The distributed detection for this type of observations can be described as the following binary hypotheses testing problem:

$$\begin{aligned} H_0^{(k)} &: X_i^{(k)} \sim [f_{k,0}(X_i^{(k)}), f_{k,0}^{(j)}(X_i^{(k)}, X_{i+j}^{(k)})] \\ H_1^{(k)} &: X_i^{(k)} \sim [f_{k,1}(X_i^{(k)}), f_{k,1}^{(j)}(X_i^{(k)}, X_{i+j}^{(k)})], \quad i = 1, \cdots, n; \; k = 1, 2 \end{aligned} \tag{69}$$

where $f_l^{(k)}(\cdot)$ is the univariate density of the observations for sensor k under H_l $(l = 0, 1)$ and the two observation sequences $\{X_i^{(k)}\}_{i=1}^N$ $(k = 1, 2)$ are modeled by the m-dependent, the ϕ-mixing or the ρ-mixing sequence.

Although the use of the Chernoff bound simplifies the analysis, the determination of the optimal thresholds involves complex computations and results in nonlinear optimization for the optimal nonlinearities. We pose additional conditions for the relationship between η_k and $(\mu_{k,i}, \sigma_{k,i}^2)$ $(i = 0, 1)$ and express η_k $(k = 1, 2)$ as functions of $(\mu_{k,i}, \sigma_{k,i}^2)$. Under such conditions for our design criteria, a large sample size is used to guarantee high-quality performance.

IV.A. Dependence Across Time

In this case, we assume that the two observation sequences are independent across sensors conditioned on either hypothesis; thus there is no correlation between the two statistics $T_{n,1}$ and $T_{n,2}$, i.e., $\rho_l = 0$ for $l = 0, 1$. By using the approximation (40), we have, for this case,

$$\begin{aligned} B_u \approx\; & pA_1 \exp\left\{-A_2 n \left[\frac{(\eta_1 - \mu_{1,0}(g_1))^2}{\sigma_{1,0}^2(g_1)} + \frac{(\eta_2 - \mu_{2,0}(g_2))^2}{\sigma_{2,0}^2(g_2)}\right]\right\} \\ & + (1-p)A_1 \exp\left\{-A_2 n \left[\frac{(\eta_1 - \mu_{1,1}(g_1))^2}{\sigma_{1,1}^2(g_1)} + \frac{(\eta_2 - \mu_{2,1}(g_2))^2}{\sigma_{2,1}^2(g_2)}\right]\right\} \end{aligned} \tag{70}$$

where A_1 and A_2 are known constants.

For the same reason as in the previous section, we pose the balance condition

$$\frac{(\eta_k - \mu_{k,0}(g_k))^2}{\sigma^2_{k,0}(g_k)} = \frac{(\eta_k - \mu_{k,1}(g_k))^2}{\sigma^2_{k,1}(g_k)} \tag{71}$$

for each sensor k. Under the assumption $\mu_{k,1}(g_k) \geq \mu_{k,0}(g_k)$, the solutions of (71) for (η_1, η_2) are

$$\eta_k = \frac{\sigma_{k,1}(g_k)\mu_{k,0}(g_k) + \sigma_{k,0}(g_k)\mu_{k,1}(g_k)}{\sigma_{k,1}(g_k) + \sigma_{k,0}(g_k)} \tag{72}$$

for $k = 1, 2$. The conditions (18)-(21) and (28) are automatically satisfied. Although these thresholds are suboptimal, they depend only on $(\mu_{k,i}, \sigma_{k,i})$, ($k = 1, 2$ and $i = 0, 1$), which can be calculated directly from the observations. Consequently, the bound given by (70) takes the form

$$\begin{aligned} B_u &= A_1 \exp\{-A_2 N[J_e(g_1, g_2)]\} \\ &= A_1 \exp\left\{-A_2 N\left[\frac{(\mu_{1,1}(g_1) - \mu_{1,0}(g_1))^2}{(\sigma_{1,1}(g_1) + \sigma_{1,0}(g_1))^2} + \frac{(\mu_{2,1}(g_2) - \mu_{2,0}(g_2))^2}{(\sigma_{2,1}(g_2) + \sigma_{2,0}(g_2))^2}\right]\right\} \end{aligned} \tag{73}$$

where A_1 and A_2 are constants.

The minimization of the above objective function (i.e., the maximization of J_e) with respect to g_k, for $k = 1, 2$, results in nonlinear integral equations, which are difficult to solve. To simplify the design procedure further, we approximate J_e in (73) as follows

$$J_e(g_1, g_2) \approx \frac{(\mu_{1,1}(g_1) - \mu_{1,0}(g_1))^2}{\gamma_1\sigma^2_{1,1}(g_1) + (1-\gamma_1)\sigma^2_{1,0}(g_1)} + \frac{(\mu_{2,1}(g_2) - \mu_{2,0}(g_2))^2}{\gamma_2\sigma^2_{2,1}(g_2) + (1-\gamma_2)\sigma^2_{2,0}(g_2)} = J_\gamma(g_1, g_2, \gamma_1, \gamma_2) \tag{74}$$

where the denominators of the two terms for the two sensors are weighted sums of $\sigma^2_{k,1}$ and $\sigma^2_{k,0}$, for $k = 1, 2$, with weights γ_k. This form of design criterion has been useful for single-sensor detection (see [12]). The weights γ_k are chosen to maximize $J(g^*_{1,\gamma_1}, g^*_{2,\gamma_2})$, where g^*_{k,γ_k} ($k = 1, 2$) are the optimum nonlinearities for fixed γ_k.

Since there is no correlation between the two sensors, maximizing $J_\gamma(g_1, g_2)$ (here the arguments γ_k are omitted) with respect to (g_1, g_2) for fixed γ_k is equivalent to maximizing

$$J_k(g_k) = \frac{(\mu_{k,1}(g_k) - \mu_{k,0}(g_k))^2}{\gamma_k \sigma^2_{k,1}(g_k) + (1-\gamma_k)\sigma^2_{k,0}(g_k)} \tag{75}$$

with respect to g_k $(k = 1, 2)$, respectively. We notice that J_k $(k = 1, 2)$ are invariant under the scaling of g_k, i.e., $J_k(g_k) = J_k(\alpha_k g_k)$ with the scaling factors α_k, and thus we can use the same technique as in the previous section (see [9]) for the optimizations of J_k. The optimizations of J_k for each sensor is equivalent to the maximizations of the objective function

$$H_k(g_k) = \mu_{k,1}(g_k) - \mu_{k,0}(g_k) + \lambda_k[\gamma_k \sigma^2_{k,1}(g_k) + (1-\gamma_k)\sigma^2_{k,1}(g_k)] \tag{76}$$

with respect to g_k, where λ_k is the Lagrange multiplier.

Performing the optimization, for $k = 1, 2$, we obtain the following linear integral equations

$$f_{k,1}(x) - f_{k,0}(x) - \int K_k(x,y) g_k(y) dy = g_k(x)[\gamma_k f_{k,1}(x) + (1-\gamma_k) f_{k,0}(x)] \tag{77}$$

where

$$\begin{aligned} K_k(x,y) \;=\; & \sum_{j=1}^{m} \gamma_k (f^{(k)}_{1,j}(x,y) - 2 f_{k,1}(x) f_{k,1}(y)) + (1-\gamma_k)(f^{(k)}_{0,j}(x,y) - 2 f_{k,0}(x) f_{k,0}(y))] \\ & - \gamma_k f_{k,1}(x) f_{k,1}(y) - (1-\gamma_k) f_{k,0}(x) f_{k,0}(y) \end{aligned}$$

is the kernel of integration with the notation

$$f^{(k)}_{i,j}(x,y) = f^{(j)}_{k,i}(X^{(k)}_1 = x, X^{(k)}_{j+1} = y) + f^{(k)}_{k,i}(X^{(k)}_1 = y, X^{(k)}_{j+1} = x). \tag{78}$$

Solving each integral equations for fixed γ_k and $k = 1, 2$ we obtain the optimum nonlinearity, which is a function of γ_k. It is straightforward to show that

$$\gamma_k \sigma^2_{k,1}(g_k) + (1-\gamma_k)\sigma^2_{k,0}(g_k) = \gamma_k var_{k,1}(g_k) + (1-\gamma_k) var_{k,0}(g_k) + \int\int \bar{K}_k(x,y) g_k(x) g_k(y) dx dy \tag{79}$$

where

$$\bar{K}_k(x,y) = K_k(x,y) + \gamma_k f_{k,1}(x) f_{k,1}(y) + (1-\gamma_k) f_{k,0}(x) f_{k,0}(y) \tag{80}$$

for $k = 1, 2$, and thus for g^*_{k,γ_k} which are the solutions of (77)

$$\gamma_k \sigma^2_{k,1}(g^*_{k,\gamma_k}) + (1-\gamma_k)\sigma^2_{k,0}(g^*_{k,\gamma_k}) = \int (f_{k,1}(x) - f_{k,0}(x)) g^*_{k,\gamma_k}(x) dx. \tag{81}$$

Finally, the optimal γ_k $(k = 1, 2)$ are obtained by maximizing

$$J_k(g^*_{k,\gamma_k}) = \int (f_{k,1}(x) - f_{k,0}(x)) g^*_{k,\gamma_k}(x) dx. \tag{82}$$

IV.B. Dependence Across Time and Sensors

We consider the case for which the observations are dependent across time and sensors. We are particularly interested in symmetric case similar to the one in Subsection III.B, i.e., $f_{1,i}(x) = f_{2,i}(x) = f_i(x)$

$$f^{(j)}_{1,i}(X^{(1)}_1, X^{(1)}_{j+1}) = f^{(j)}_{2,i}(X^{(2)}_1, X^{(2)}_{j+1}) = f^{(j)}_i(X_1, X_{j+1}) \tag{83}$$

conditioned on H_i $(i = 0, 1)$ and $\eta_1 = \eta_2 = \eta$. Again, the same form of the nonlinearity $g_1(\cdot) = g_2(\cdot) = g(\cdot)$ is used. Therefore, $\mu_{1,i} = \mu_{2,i} = \mu_i$ and $\sigma_{1,i} = \sigma_{2,i} = \sigma_i$, for $i = 0, 1$. In the present case, the upper bound given by (29) takes the form

$$\begin{aligned} B_u \;=\; & \sqrt{2\pi} p \left\{ 2c_1 \exp\left[\frac{-n(\eta-\mu_0)^2}{2\sigma_0^2}\right] + c_2 \exp\left[\frac{-n(\eta-\mu_0)^2}{(1+\rho_0)\sigma_0^2}\right] \right\} \\ & + \sqrt{2\pi}(1-p) \left\{ 2c_1 \exp\left[\frac{-n(\eta-\mu_1)^2}{2\sigma_1^2}\right] + c_2 \exp\left[\frac{-n(\eta-\mu_1)^2}{(1+\rho_1)\sigma_1^2}\right] \right\} \end{aligned} \tag{84}$$

which consists of two parts from the two hypotheses. Instead of optimizing the above bound directly, we proceed in an *ad hoc* way to determine the nonlinearity. Basically there are two different forms of exponents in B_u, namely

$$B_{u_1} == \sqrt{2\pi} p 2c_1 \exp[\frac{-n(\eta-\mu_0)^2}{2\sigma_0^2}] + \sqrt{2\pi}(1-p) 2c_1 \exp[\frac{-n(\eta-\mu_1)^2}{2\sigma_1^2}] \tag{85}$$

and

$$B_{u_2} = \sqrt{2\pi}pc_2 \exp[\frac{-n(\eta-\mu_0)^2}{(1+\rho_0)\sigma_0^2}] + c_2 \exp[\frac{-n(\eta-\mu_1)^2}{(1+\rho_1)\sigma_1^2}] \tag{86}$$

As discussed in the case of a weak signal, we first simplify the problem by posing the balance conditions

$$\frac{(\eta-\mu_0(g))^2}{2\sigma_0^2(g)} = \frac{(\mu_1(g)-\eta)^2}{2\sigma_1^2(g)} \tag{87}$$

for B_{u_1} and

$$\frac{(\eta-\mu_0(g))^2}{(1+\rho_0(g))\sigma_0^2(g)} = \frac{(\mu_1(g)-\eta)^2}{(1+\rho_1(g))\sigma_1^2(g)} \tag{88}$$

for B_{u_2}, such that the threshold can be set for the design criterion. The solutions of (87) and (88) are

$$\eta = \frac{\sigma_1(g)\mu_0(g) + \sigma_0(g)\mu_1(g)}{\sigma_1(g)+\sigma_0(g)} \tag{89}$$

and

$$\eta = \frac{(1+\rho_1(g))^{\frac{1}{2}}\sigma_1(g)\mu_0(g) + (1+\rho_0(g))^{\frac{1}{2}}\sigma_0(g)\mu_1(g)}{(1+\rho_1(g))^{\frac{1}{2}}\sigma_1(g) + (1+\rho_0(g))^{\frac{1}{2}}\sigma_0(g)} \tag{90}$$

respectively. Again conditions (18)-(21) and (28) are satisfied with both choices of threshold. We notice that, by using the above thresholds, the exponents in B_{u_1} and B_{u_2} essentially take the form

$$-\frac{(\mu_1(g)-\mu_0(g))^2}{(\sigma_1(g)+\sigma_0(g))^2} \tag{91}$$

and

$$-\frac{(\mu_1(g)-\mu_0(g))^2}{[(1+\rho_1(g))^{\frac{1}{2}}\sigma_1(g) + (1+\rho_0(g))^{\frac{1}{2}}\sigma_0(g)]^2} \tag{92}$$

Since the optimization of either (91) or (92) with respect to $g(\cdot)$ leads to nonlinear integral equations, we use approximations similar to the one in Subsection IV.A In summary, the criterion for optimum nonlinearity $g(\cdot)$ are

$$J_a(g) = \frac{(\mu_1(g)-\mu_0(g))^2}{\gamma_a\sigma_1^2(g) + (1-\gamma_a)\sigma_0^2(g)} \tag{93}$$

and

$$J_b(g) = \frac{(\mu_1(g) - \mu_0(g))^2}{\gamma_b(1+\rho_1(g))\sigma_1^2(g) + (1-\gamma_b)(1+\rho_0(g))\sigma_0^2(g)}. \tag{94}$$

Then, for fixed γ_a and γ_b, we maximize J_a and J_b with respect to $g(\cdot)$ independently to obtain two different forms nonlinearities $g_a(\cdot)$ and $g_b(\cdot)$, respectively; the respective γ_a and γ_b are chosen by maximizing $J(g^*_{a,\gamma_a})$ and $J(g^*_{b,\gamma_b})$. Finally, the nonlinearity employed by the two sensors is

$$g^*(x) = \tau_s g_a^*(x) + (1-\tau_s) g_b^*(x) \tag{95}$$

where the parameter τ_s is set by numerical tests.

The maximization of (93) with respect to g and γ_a has been discussed in Subsection IV.A. The linear integral equation satisfied by the optimum g_a^* is

$$f_1(x) - f_0(x) - \int K(x,y) g_a^*(y) dy = g_a^*(x)[\gamma_a f_1(x) + (1-\gamma_a) f_0(x)] \tag{96}$$

where

$$\begin{aligned} K(x,y) \;=\; & \sum_{j=1}^{m} \{\gamma_a [f_{1,j}(x,y) - 2f_1(x)f_1(y)] + (1-\gamma_a)[f_{0,j}(x,y) - 2f_0(x)f_0(y)]\} \\ & -\gamma_a f_1(x) f_1(y) - (1-\gamma_a) f_0(x) f_0(y) \end{aligned} \tag{97}$$

is the kernel of integration and

$$f_{i,j}(x,y) = f_i(X_1 = x, X_{j+1} = y) + f_i(X_1 = y, X_{j+1} = x). \tag{98}$$

Because J_b is also invariant under the scaling of g [$J_b(sg) = J_b(g)$ for constant s], equivalently, we maximize the following form

$$H_b(g) = \mu_1(g) - \mu_0(g) + \lambda[\gamma_b(1+\rho_1(g))\sigma_e^2(g) + (1-\gamma_b)(1+\rho_0(g))\sigma_0^2(g)] \tag{99}$$

with respect to g for fixed γ_b, where λ is the Lagrange multiplier. Using the same technique as in Subsection IV.A, we obtain the following linear integral equation satisfied by the optimum g_b^*

$$f_1(x) - f_0(x) - \int [K_c(x,y) + \tilde{K}_c(x,y)] g_b^*(y) dy = [\gamma_b f_1(x) + (1-\gamma_b) f_0(x)] g_b^*(x) \tag{100}$$

where

$$K_c(x,y) = 2\sum_{j=1}^{m}\{\gamma_b[f_1^{(j)}(x,y) - f_1(x)f_1(y)] + (1-\gamma_b)[f_0^{(j)}(x,y) - f_0(x)f_0(y)]\}$$
$$-\gamma_b f_1(x)f_1(y) - (1-\gamma_b)f_0(x)f_0(y) \quad (101)$$

is the kernel of integration and

$$\tilde{K}_c(x,y) = \gamma_b[\tilde{f}_1^{(0)}(x,y) - f_1(x)f_1(y)] + (1-\gamma_b)[\tilde{f}_0^{(0)}(x,y) - f_0(x)f_0(y)]$$
$$+2\sum_{j=1}^{m}\{\gamma_b[\tilde{f}_1^{(j)}(x,y) - f_1(x)f_1(y)] + (1-\gamma_b)[\tilde{f}_0^{(j)}(x,y) - f_0(x)f_0(y)]\} \quad (102)$$

with $\tilde{f}_i^{(j)}(\cdot,\cdot)$ the joint density of $X_1^{(1)}$ and $X_{j+1}^{(2)}$ across sensors.

V. NUMERICAL RESULTS

To support the analysis for the proposed scheme in previous sections, we consider the ρ-mixing Cauchy noise for the weak signal model characterized in Section III, and the observations of ρ-mixing Rayleigh densities (under H_0) vs. the ones of ρ-mixing log-normal densities (under H_1) for the strong signal model characterized in Section IV. The ms are truncated to values which are large enough for the simulations in different examples. The sample size of the simulations is 100000 for weak signals and 1000 for strong signals.

V.A. Weak Signals

We first consider a weak signal in additive Cauchy noise which is symmetric with respect to its medium. The performance by using the optimal nonlinearities is compared with the one by ignoring the dependence across time and/or sensors for the two cases given in Section III.

Without loss of generality, the the weak signal θ is assumed to be 0.06 for convenience. The univariate density of the Cauchy noise in our numerical examples is given by

$$f_k(x^{(k)}) = \frac{1}{\pi[1+(x^{(k)}-\gamma_k)^2]}, \quad k = 1, 2$$

where $-\infty < \gamma_k < \infty$ is the median (assumed known).

Although the second-order joint density of a Cauchy noise is difficult to characterize directly, it can be calculated from the second-order joint density of a Gaussian process by a nonlinear transformation (see [6]). Let

$$f_G(x^{(k)}) = \frac{\exp[-(x^{(k)}-\gamma_k)^2/2]}{\sqrt{2\pi}}$$

and

$$f_G(x^{(k)}, y^{(k)}) = \frac{1}{2\pi(1-\rho_G^{(k)})} \exp\{\frac{-1}{2(1-\rho_G^{(k)})}[(x^{(k)}-\gamma_k)^2+(y^{(k)}-\gamma_k)^2-2\rho_G^{(k)}(x^{(k)}-\gamma_k)(y^{(k)}-\gamma_k)]\}$$

be the univariate and the second-order joint densities of the underlying Gaussian process, where $\rho_G^{(k)}$ is the correlation coefficient. Then the nonlinear transformation mentioned above has the form $T(x) = \tan[\pi \cdot erf(x/\sqrt{2})/2]$ (see [9]). In addition, we assume that the underlining Gaussian process for the corresponding Cauchy noise of each sensor is characterized the following autoregressive model with the correlation coefficients $-1 \leq \rho_k \leq 1$ for $k = 1, 2$

$$\begin{aligned} N_1^{(k)} &= V_1^{(k)} \\ N_i^{(k)} &= \rho_k N_{i-1}^{(k)} + \sqrt{1-\rho_k^2} V_i^{(k)};\ i > 1 \end{aligned}$$

where both $V_i^{(k)}$, for $i = 1, 2, \cdots, n$; $k = 1, 2$, are sequences of *i.i.d.* Gaussian random variables and have standard Gaussian densities. In the following examples, for the cases of dependence across time only (Examples 1 and 3), $V_i^{(1)}$, $V_i^{(2)}$ for $i = 1, 2, \cdots, n$ are generated independently of each other; for the cases of dependence across time and sensors (Examples 2 and 4), they are generated dependently as follows

$$V_i^{(2)} = \rho_c V_i^{(1)} + \sqrt{1-\rho_c^2} W_i$$

where $V_i^{(1)}$ and W_i are two independent *i.i.d.* Gaussian processes and $-1 \leq \rho_c \leq 1$ is another correlation parameter.

In the following examples, we give the plots of the optimal nonlinearities and figures of the receiver operating characteristic (ROC) type for the above weak signal in additive noise. In the figures of ROC, the two types of average cost, E_0 and E_1, are normalized with respect to p and $1 - p$, respectively.

Example 1: This example is for the case of dependence across time only. The parameters of the above noise model are set to be $\gamma_1 = 0.0$, $\rho_1 = 0.95$ and $m = 150$ for sensor 1; $\gamma_2 = 0.50$, $\rho_2 = 0.90$ and $m = 100$ for sensor 2. The costs are chosen as $c_1 = c_1' = 1.0$ and $c_2 = c_2' =$

0.50. The optimal nonlinearities $g_{k,opt}(x)$ $(k = 1,2)$, and the ones obtained by ignoring the dependence $g_{k,iid}(x) = -f_k'(x)/f_k(x)$ $(k = 1,2)$, are given in Figures 1.1 and 1.2 for the two sensors, respectively. The average costs resulted by employing these two forms of nonlinearities are shown in Figure 1.3 for sensor 1 under two fixed thresholds of sensor 2 which give the best performances of sensor 2 (at their individual crossing points in Figure 1.4 for two distinct ROCs), and in Figure 1.4 for sensor 2 for two fixed thresholds of sensor 1 which give the best performance of sensor 1 (at their individual crossing points in Figures 1.3 for two distinct ROCs). In Figures 1.3 and 1.4 the solid lines represent the ROCs of employing the $g_{k,opt}$ $(k = 1,2)$ and the dot-dash lines represent the ROCs of employing the $g_{k,iid}$ $(k = 1,2)$. From Figures 1.3 and 1.4 we conclude that the performance obtained by employing $g_{k,iid}$ is better than the one obtained by ignoring the dependence across time for this specific example.

Example 2: This example is for the scheme in the case of Subsection III.B. Here we set $\gamma_1 = \gamma_2 = 0.0$, $\rho_1 = \rho_2 = 0.95$ and $m = 200$ for both sensors, and the correlation parameter across time is set to be $\rho_c = -0.90$. After numerical tests, we obtain $\tau_w = 0.0225$ in Eq. (57). In Figure 2.1, we draw the nonlinearity g_b^* given in Eq. (57), which is used by both sensors. In that figure, the optimal nonlinearities obtained by ignoring the dependence across sensors g_{idd} (represented by dash lines), as well as across time and sensors g_{iid} (represented by dot-dash lines) are also included. Notice that $g_a^*(x) = g_{idd}(x)$ in Eq. (57). In Figure 2.2, we give the plots of the average costs E_0/p and $E_1/(1-p)$ (ROC) as a function of the threshold used by both sensors. The solid lines represent the ROC of employing the $g_{opt} = \tau_w g_{idd} + (1-\tau_w)g_b^*$, the dash lines represent the ROC of employing the g_{idd}, and the dot-dash lines represent the ROC of employing the g_{iid}. From Figure 2.2 we conclude the performance obtained by employing g_{opt} is the best for this specific example.

V.B. Strong Signals

Similarly, we now show the performance for the strong signal model described in Section IV. The two observations hypotheses under considerations are characterized by the stationary log-normal densities (univariate and second-order joint) under H_1, and the stationary Rayleigh densities (univariate and second-order joint) under H_0. The log-normal processes $X_{1,i}^{(k)}$ $(i = 1, 2, \cdots; k = 1, 2)$ are obtained from the nonlinear transformations $X_{1,i}^{(k)} = \exp[N_{1,i}^{(k)} + \mu_{x,1}^{(k)}]$ in which $N_{1,i}^{(k)}$ are generated as follows

$$\begin{aligned} N_{1,1}^{(k)} &= V_{1,1}^{(k)} \\ N_{1,i}^{(k)} &= \rho_{1,k} N_{1,i-1}^{(k)} + \sqrt{1-\rho_{1,k}^2}\, V_{1,i}^{(k)};\ i > 1 \end{aligned}$$

where $\rho_{1,k}$ are correlation coefficients and $V_{1,i}^{(k)}$ $(i = 1, 2, \cdots, n)$ for sensor k $(k = 1, 2)$ is a sequence of *i.i.d.* Gaussian random variables and has Gaussian density $\mathcal{N}(\prime, (\sigma_{\sqsubseteq,\infty}^{(\|)})^{\in})$. In the case of dependence across time only $V_{1,i}^{(1)}$ and $V_{1,i}^{(2)}$ are generated independently; in the case of dependence across time and sensors, they are generated dependently as

$$V_{1,i}^{(2)} = \rho_{v,1} V_{1,i}^{(1)} + \sqrt{1-\rho_{v,1}^2}\, W_{1,i}$$

where $V_{1,i}^{(1)}$ and $W_{1,i}$ are two independent *i.i.d.* Gaussian processes and $-1 \leq \rho_{v,1} \leq 1$ is an correlation parameter. The Rayleigh processes under H_0, $X_{0,i}^{(k)}$, are generated by two independent Gaussian processes $Y_{0,i}^{(k)}$ and $Z_{0,i}^{(k)}$ $(i = 1, 2, \cdots; k = 1, 2)$ as $X_{0,i}^{(k)} = \sqrt{(Y_{0,i}^{(k)})^2 + (Z_{0,i}^{(k)})^2}$ where $Y_{0,i}^{(k)}$ and $Z_{0,i}^{(k)}$ are generated as follows

$$\begin{aligned} Y_{0,1}^{(k)} &= V_{0,1}^{(k)} \\ Y_{0,i}^{(k)} &= \rho_{0,k} Y_{0,i-1}^{(k)} + \sqrt{1-\rho_{0,k}^2}\, V_{0,i}^{(k)};\ i > 1 \end{aligned}$$

and

$$Z_{0,1}^{(k)} = U_{0,1}^{(k)}$$

$$Z_{0,i}^{(k)} = \rho_{0,k} Z_{0,i-1}^{(k)} + \sqrt{1-\rho_{0,k}^2} U_{0,i}^{(k)};\ i > 1$$

where $\rho_{0,k}$ are correlation coefficients and for sensor k ($k = 1,2$) $V_{0,i}^{(k)}$, $U_{0,i}^{(k)}$ ($i = 1,2,\cdots$) are two i.i.d. Gaussian sequences with the same density $\mathcal{N}(\prime, \sigma_{\sqsubseteq,\prime}^{\in})$. In the case of dependence across time only $(V_{0,i}^{(1)}, V_{0,i}^{(2)})$ are generated independently, so is the pair $(U_{0,i}^{(1)}, U_{0,i}^{(2)})$); in the case of dependence across time and sensors, they are generated dependently as

$$V_{0,i}^{(2)} = \rho_{v,0} V_{0,i}^{(1)} + \sqrt{1-\rho_{v,0}^2} W_{0,i}$$

and

$$U_{0,i}^{(2)} = \rho_{v,0} U_{0,i}^{(1)} + \sqrt{1-\rho_{v,0}^2} W_{0,i}$$

where $W_{0,i}$ is another i.i.d. Gaussian sequence with the density $\mathcal{N}(\prime, (\sigma_{\sqsubseteq,\prime}^{(\infty)})^{\in})$ and $-1 \leq \rho_{v,0} \leq 1$ is an correlation parameter; $U_{0,i}^{(1)}$, $V_{0,i}^{(1)}$, and $W_{0,i}$ are mutually independent.

Example 3: This example is for the case of dependence across time only characterized in Subsection IV.A. The ms are set to be 300 for both sensors under the two hypotheses. For sensor 1, the parameter values are set to be $\sigma_{v,0}{}^{(1)} = 2.0$ and $\rho_{1,0} = 0.926$ under H_0 and $\sigma_{v,1}^{(1)} = 0.242$, $\mu_{x,1}^{(1)} = 0.798$ and $\rho_{1,1} = 0.992$ under H_1; for sensor 2, $\sigma_{v,0}^{(2)} = 2.20$ and $\rho_{2,0} = 0.90$ under H_0, and $\sigma_{v,1}^{(2)} = 0.242$, $\mu_{x,1}^{(2)} = 0.893$ and $\rho_{2,1} = 0.990$. After numerical tests for the optimization search, we obtain $\gamma_1 = \gamma_2 = 0.9231$ in Eqs. (75)-(77) and (79)-(82). In Figures 3.1 and 3.2, the two optimal nonlinearities $g_{k,opt}$ ($k = 1,2$) for the strong signal model are given for the two sensors; there the nonlinearities $g_{k,iid}(x) = \ln[f_{k,1}(x)/f_{k,0}(x)]$ ($k = 1,2$) are also described by dot-dash lines. The ROCs similar to the ones in Example 1 are drawn in Figures 3.3 and 3.4 for the two sensors, respectively. In Figures 3.3 and 3.4, the distinct ROCs are represented by lines similar to the ones in Example 1. The ROCs of employing $g_{k,h}$ ($k = 1,2$) which are obtained by using $\gamma_1 = \gamma_2 = 0.5$ are also included in Figures 3.3 and 3.4 (represented

by dotted lines). From the figures of ROCs we conclude that the performance obtained by employing the $g_{k,opt}$ of this strong signal model is the best for this specific example.

Example 4: This example is for the case of dependence across time and sensors characterized in Subsection IV.B. The parameter values are the same as those of sensor 1 in Example 3 except that $m = 600$. We also set $\rho_{v,0} = 0.90$ and $\rho_{v,1} = 0.99$. By using numerical tests, we set $\gamma_a = \gamma_b = 0.9231$ in Subsection IV.B, and $\tau_s = 0.95$ in Eq. (95). The nonlinearity g_b^* of this strong signal model which is used in Eq. (95) and is shown in Figure 4.1; there the nonlinerities obtained by ignoring the dependence across sensors g_{idd} (represented by dash lines), and across time and sensors g_{iid} (represented by dot-dash lines) are also included. The ROCs similar to the ones in Example 2 are drawn in Figure 4.2, where distinct ROCs are represented by lines similar to those in Example 2. Again in Figure 4.2, we include the ROCs of employing g_h which is obtained by using $\gamma_a = \gamma_b = 0.5$ (represented by dotted lines). From the ROCs we conclude that the performance obtained by employing $g_{opt} = \tau_w g_{idd} + (1 - \tau_w) g_b^*$ is better than the one obtained by ignoring the dependence across time, or across time and sensors for this specific example.

VI. CONCLUSIONS

In this paper, we extended the scheme of memoryless single-sensor detection to distributed detection for two types of signal: the weak and the strong singles which have stationary univariate and second-order densities. The model of dependence used in this paper for the correlation of noise (across time and/or sensors) is characterized by one of the mixing sequences: the m-dependent, the ϕ-mixing and the ρ-mixing sequences. The performance is measured by the average cost which couples the sensor decisions together and consists of the error probabilities involved in distributed detection. By using the two-dimensional Chernoff bounds on the error probabilities and some approximation techniques, we derived the design criteria for the memoryless nonlinearities of the sensors, which are the generalizations of the AREs in corresponding single-sensor detection. Optimization of the design criteria for the nonlinearities in different cases led to linear integral equations.

The numerical results from simulations indicate that in performance, the proposed schemes by employing optimal memoryless nonlinearities which consider the dependence (across time and/or sensors) overmatch the corresponding ones by ignoring that dependence for different cases of observation descriptions (weak signal and strong signal) and dependence (dependence across time only and dependence across time and sensors).

Although we focused on two-sensor detection, the schemes and the techniques which are used in this work can be applied to multi-sensor detection in which frequently the continuous-time optimization methods are necessary.

References

[1] R. R. Tenny and N. R. Sandell, "Detection with distributed sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-17, pp. 501-510, July 1981.

[2] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. on Aerospace and Electronic Systems* , Vol. AES-22, pp. 98-101, January 1986.

[3] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems,* vol.2, 1988.

[4] D. Teneketzis and Y. C. Ho, "The decentralized Wald problem," preprint.

[5] H. R. Hashemi and I. B. Rhodes, "Decentralized sequential detection," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 509-520, May 1989.

[6] P. Billingsley, *Convergence of Probability Measures.* New York: Wiley, 1968.

[7] R. C. Bradley, "Basic properties of strong mixing conditions ," *Dependence in Probability and Statistics*, Birkhauser, 1985.

[8] D. Sauder and E. Geraniotis, "Memoryless discrimination of mixing processes", *Proc. 1988 Conf. Inform. Sci. and Sys.*, pp. 599-604, Princeton Univ., 1988.

[9] H. V. Poor and J. B. Thomas, "Memoryless discrete-time detection of a constant signal in m-dependent noise," *IEEE Trans. Inform. Theory,* vol. IT-25, pp. 54-61, Jan. 1979.

[10] D. R. Halverson and G. L. Wise, "Discrete-time detection in $\phi-$mixing noise," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 189-198, 1980.

[11] J. S. Sadowsky, "A maximum variance model for robust detection and estimation with dependent data," *IEEE Trans. Inform. Theory,* vol. IT-32, pp. 220-226, Mar. 1986.

[12] W. A. Gardner, " A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. COM-28, pp. 807-816, June 1980.

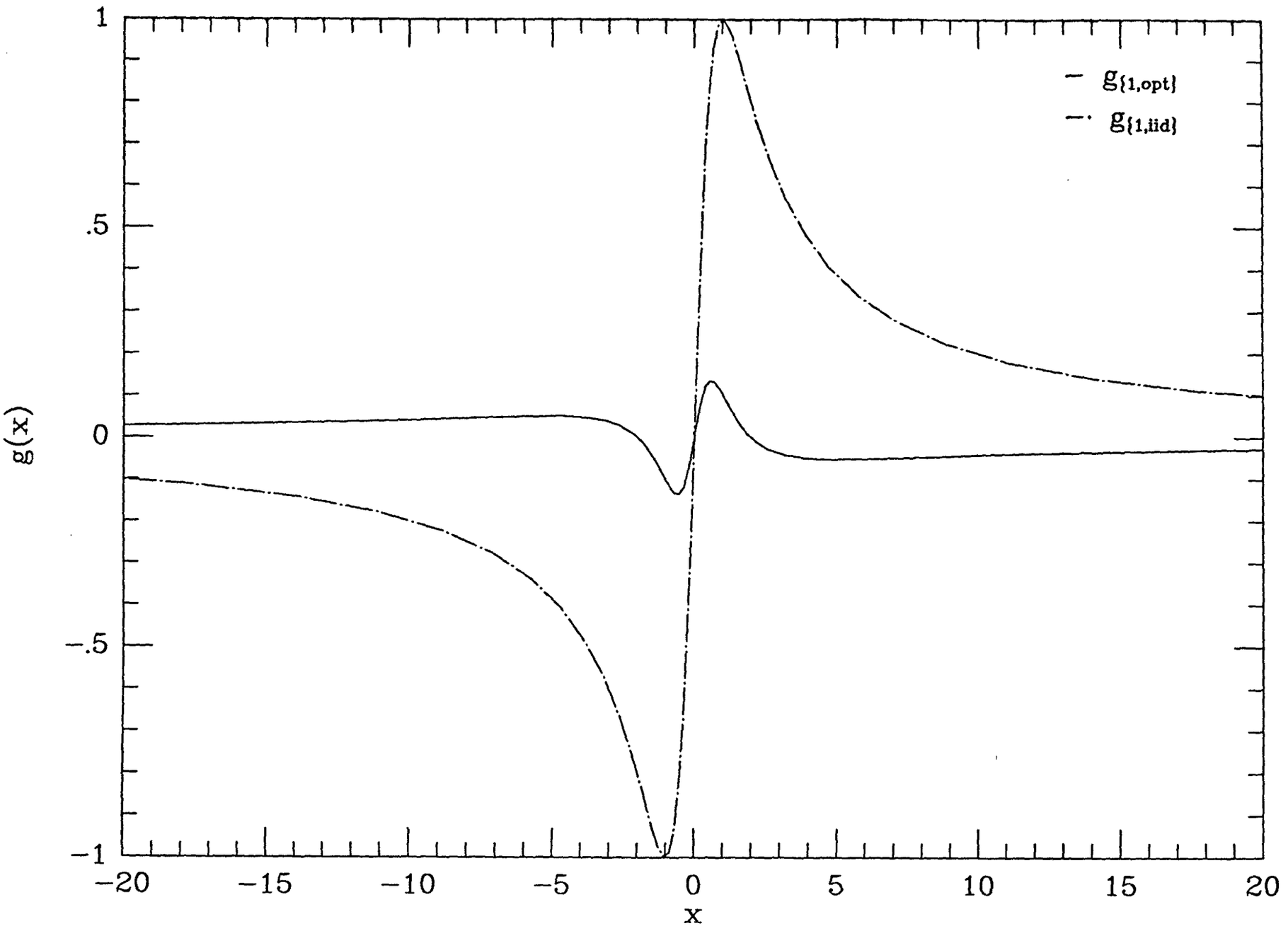

Fig. 1.1 The Nonlinearities of Sensor 1 in Example 1

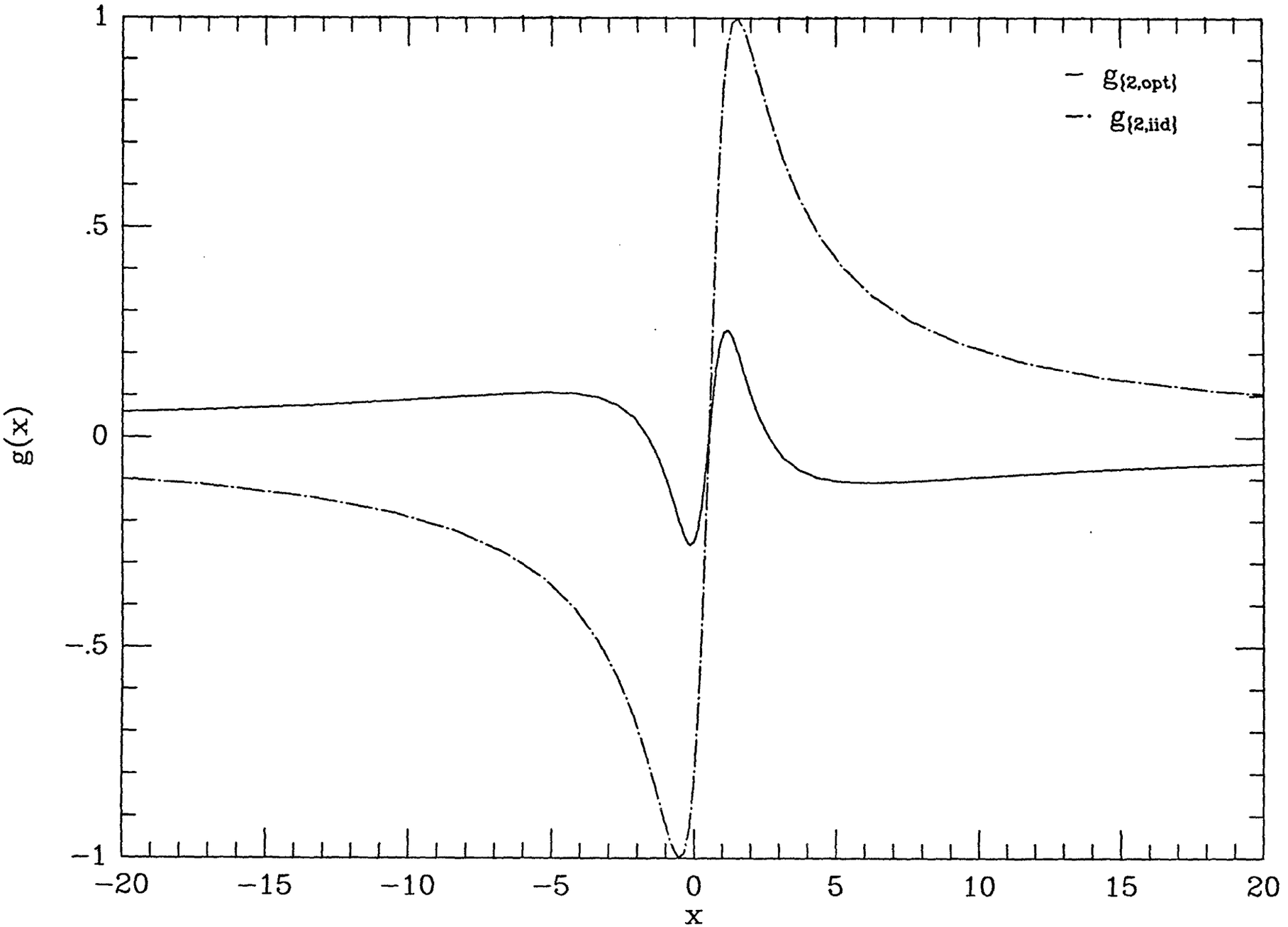

Fig. 1.2 The Nonlinearities of Sensor 2 in Example 1

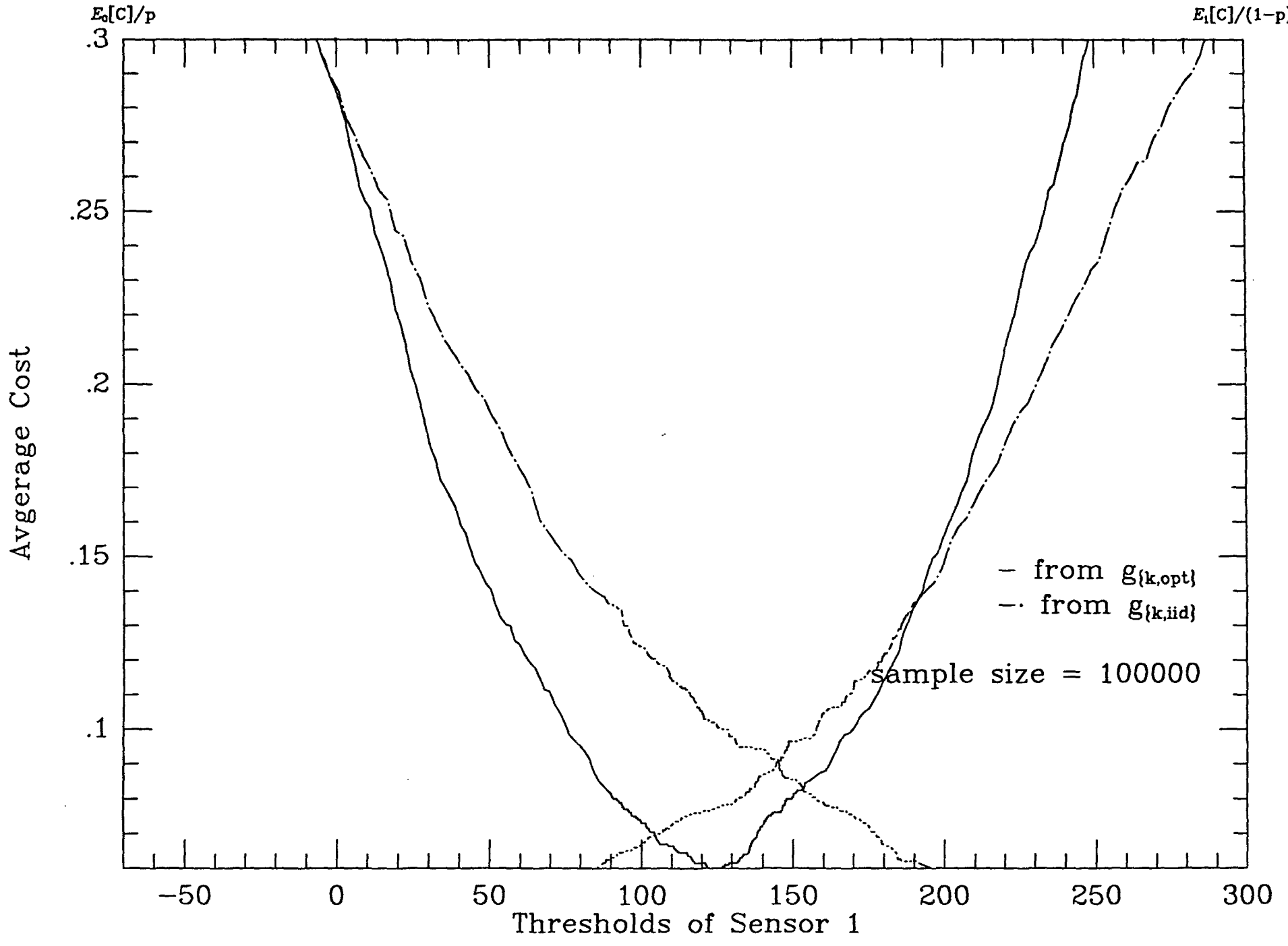

Fig. 1.3 ROCs for Example 1 under a Fixed Threshold of Sensor 2

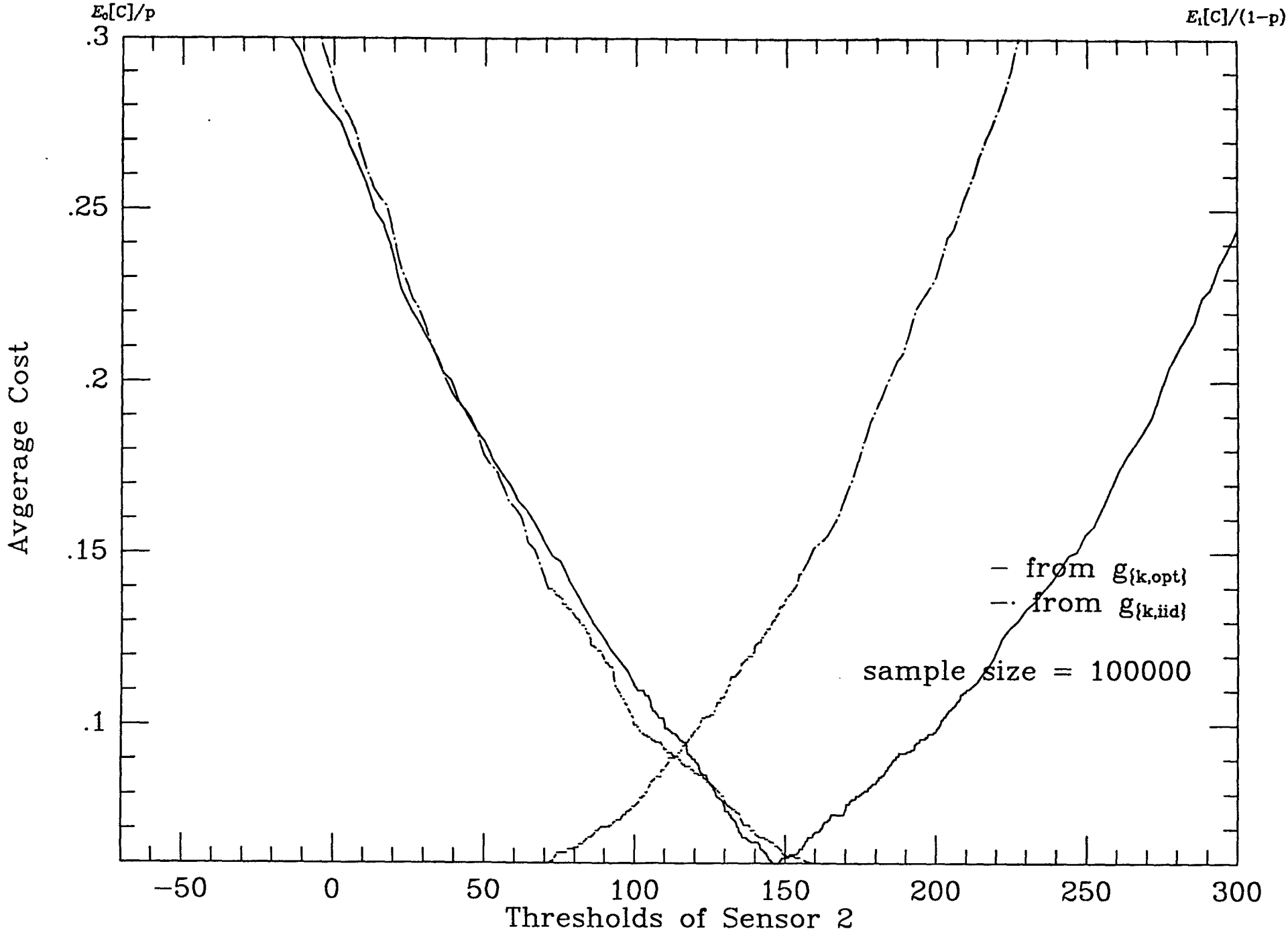

Fig. 1.4 ROCs for Example 1 under a Fixed Threshold of Sensor 1

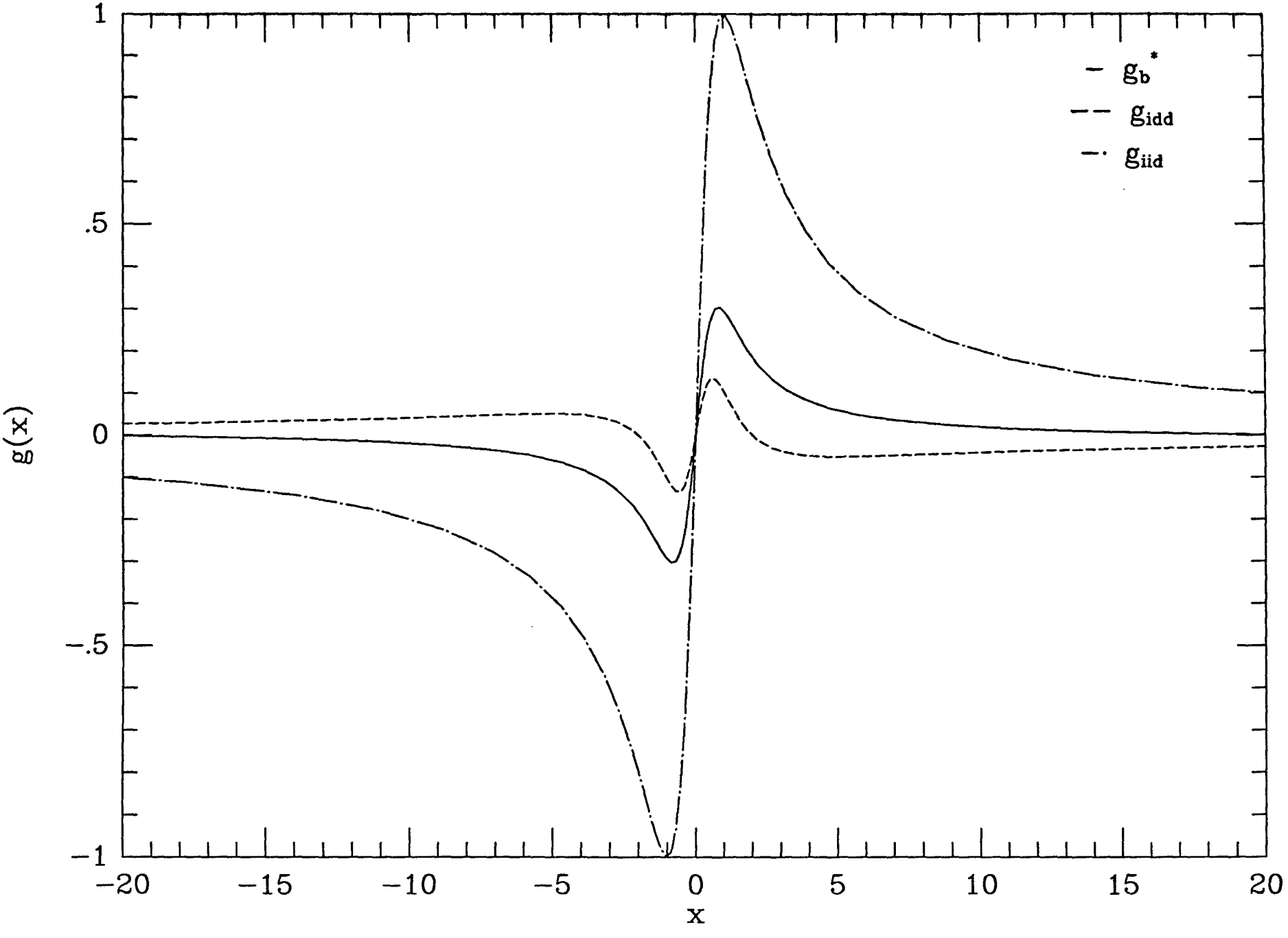

Fig. 2.1 The Nonlinearities of the Two Sensors in Example 2

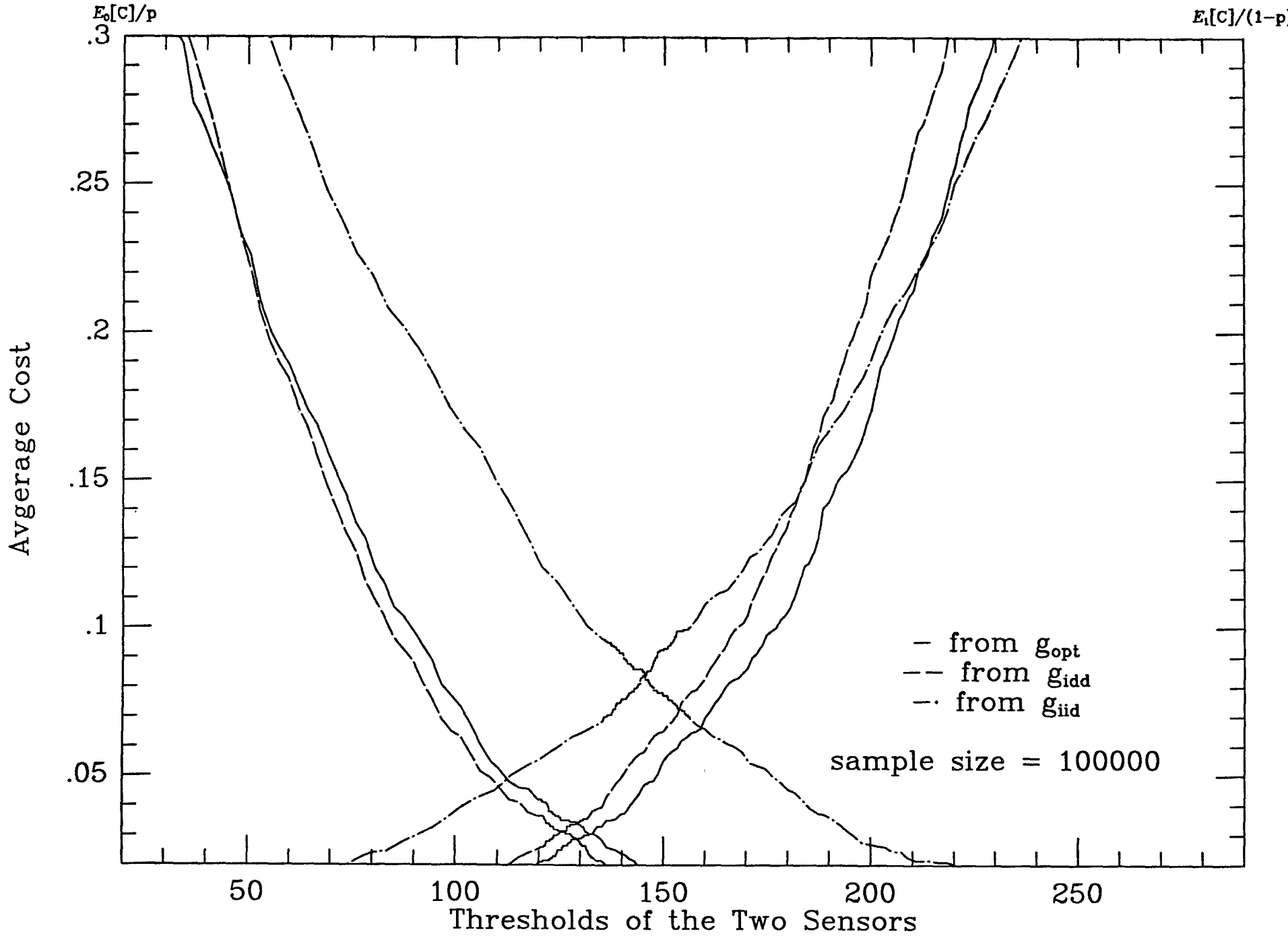

Fig. 2.2 ROCs for Example 2

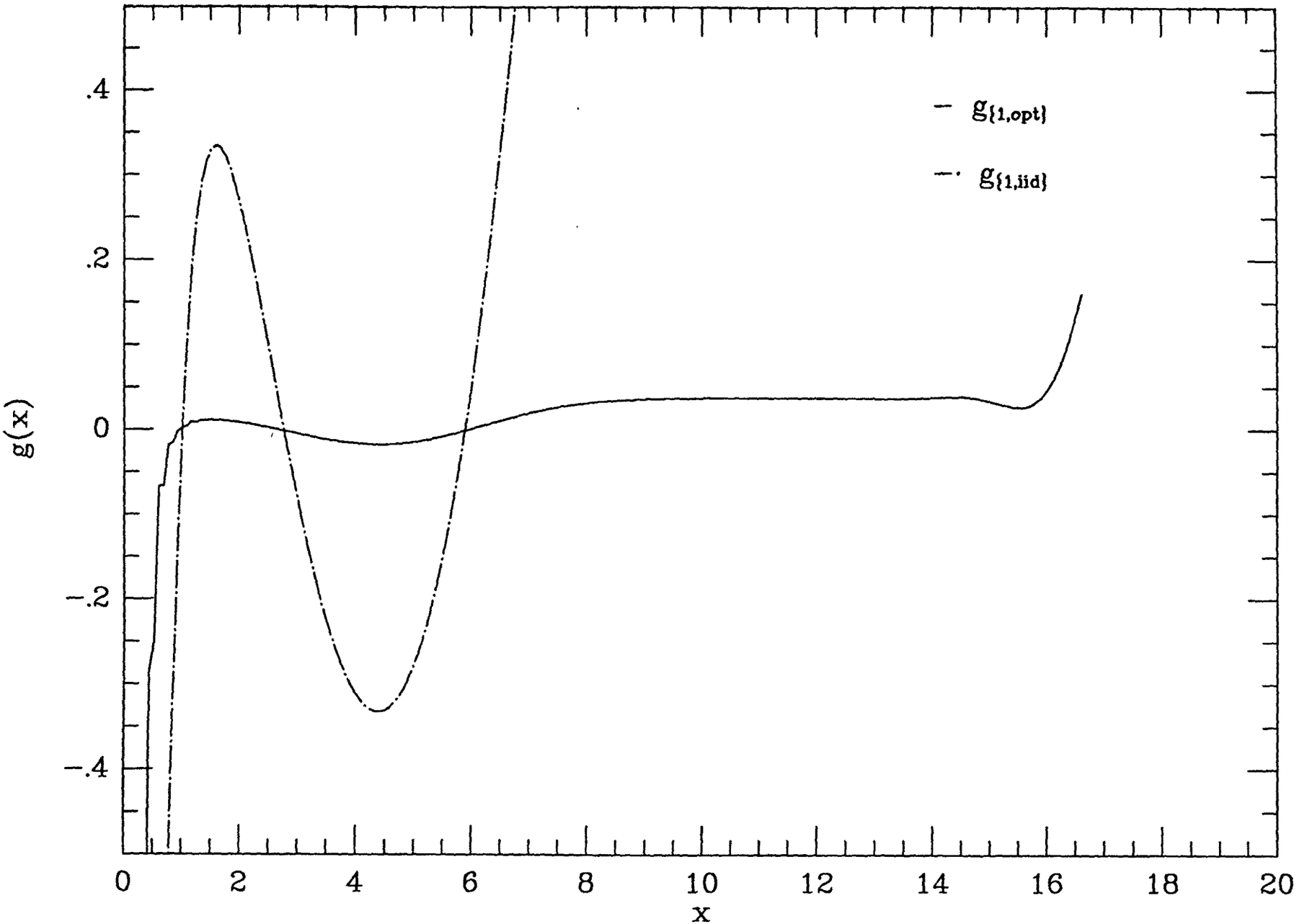

Fig. 3.1 The Nonlinearities of Sensor 1 in Example 3

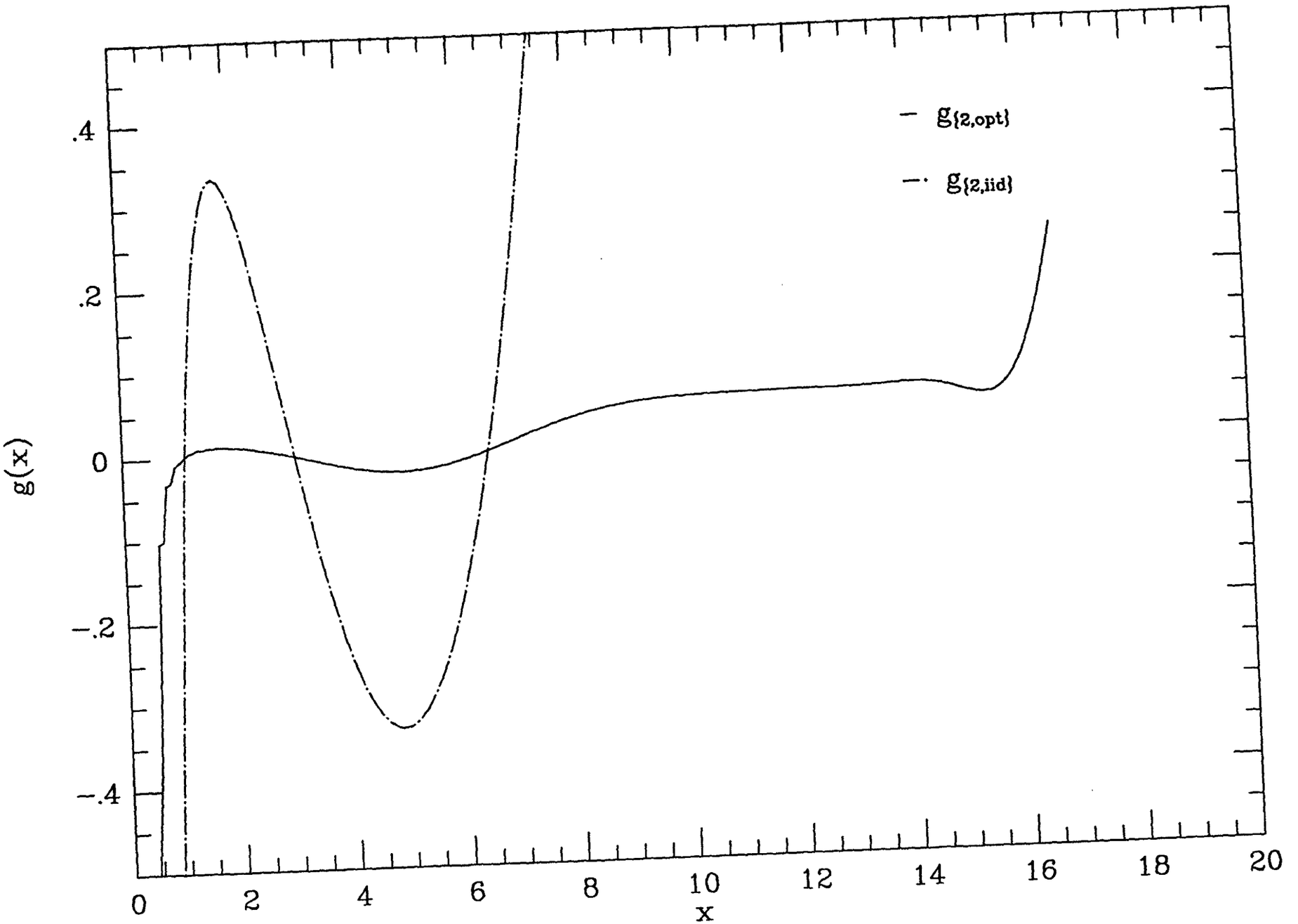

Fig. 3.2 The Nonlinearities of Sensor 2 in Example 3

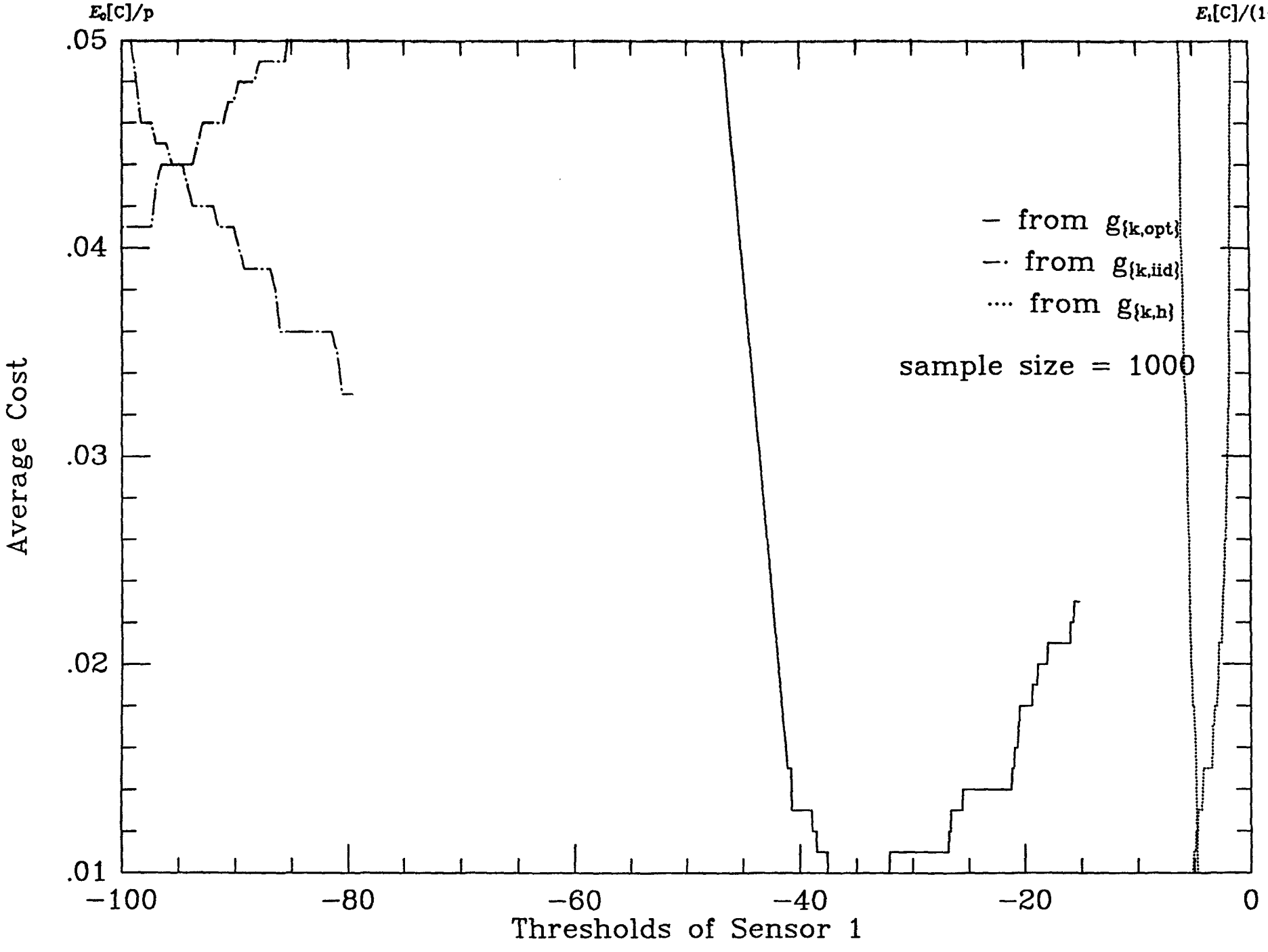

Fig. 3.3 ROCs for Example 3 under a Fixed Threshold of Sensor 2

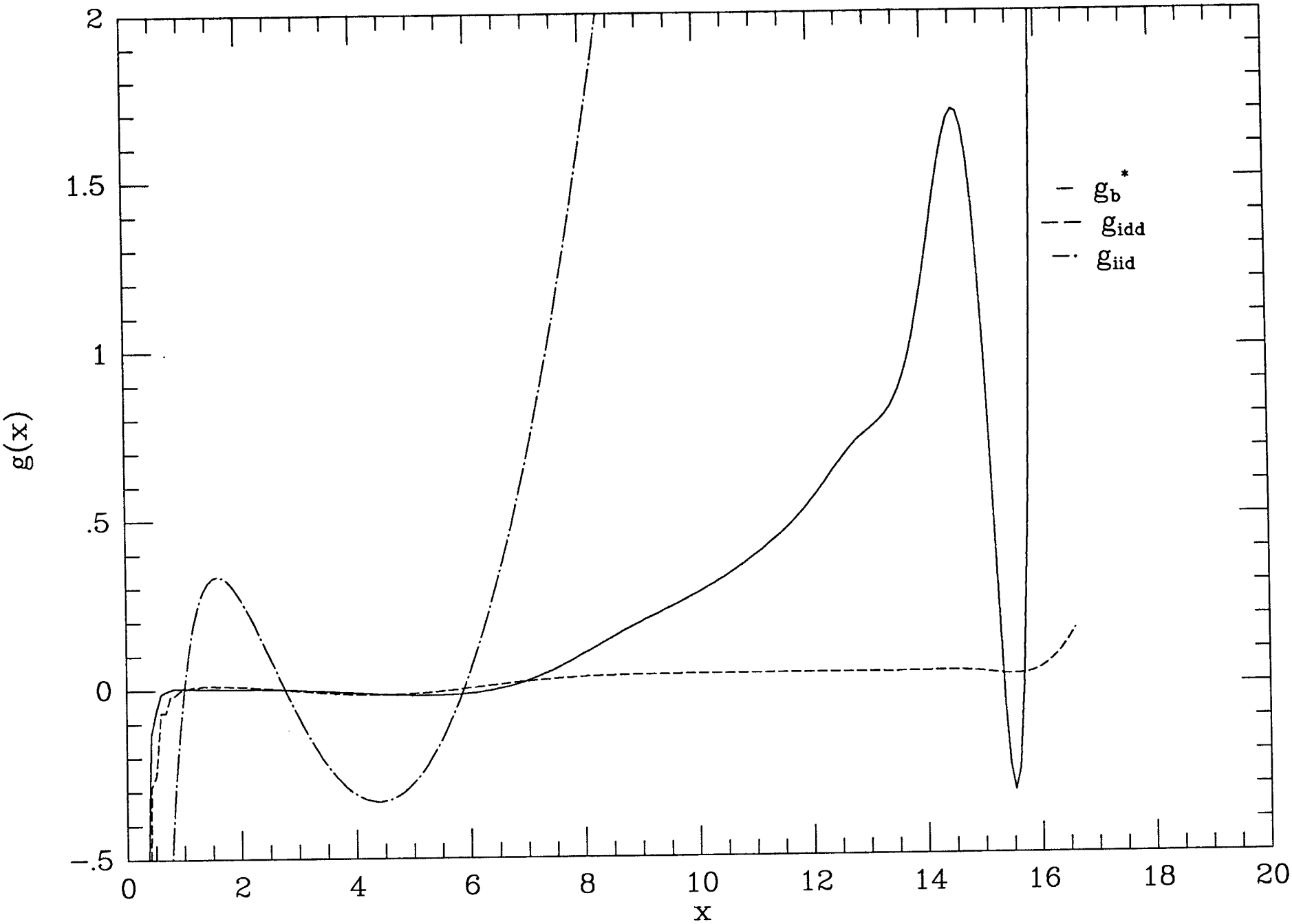

Fig. 4.1 The Nonlinearities of the Two Sensors in Example 4

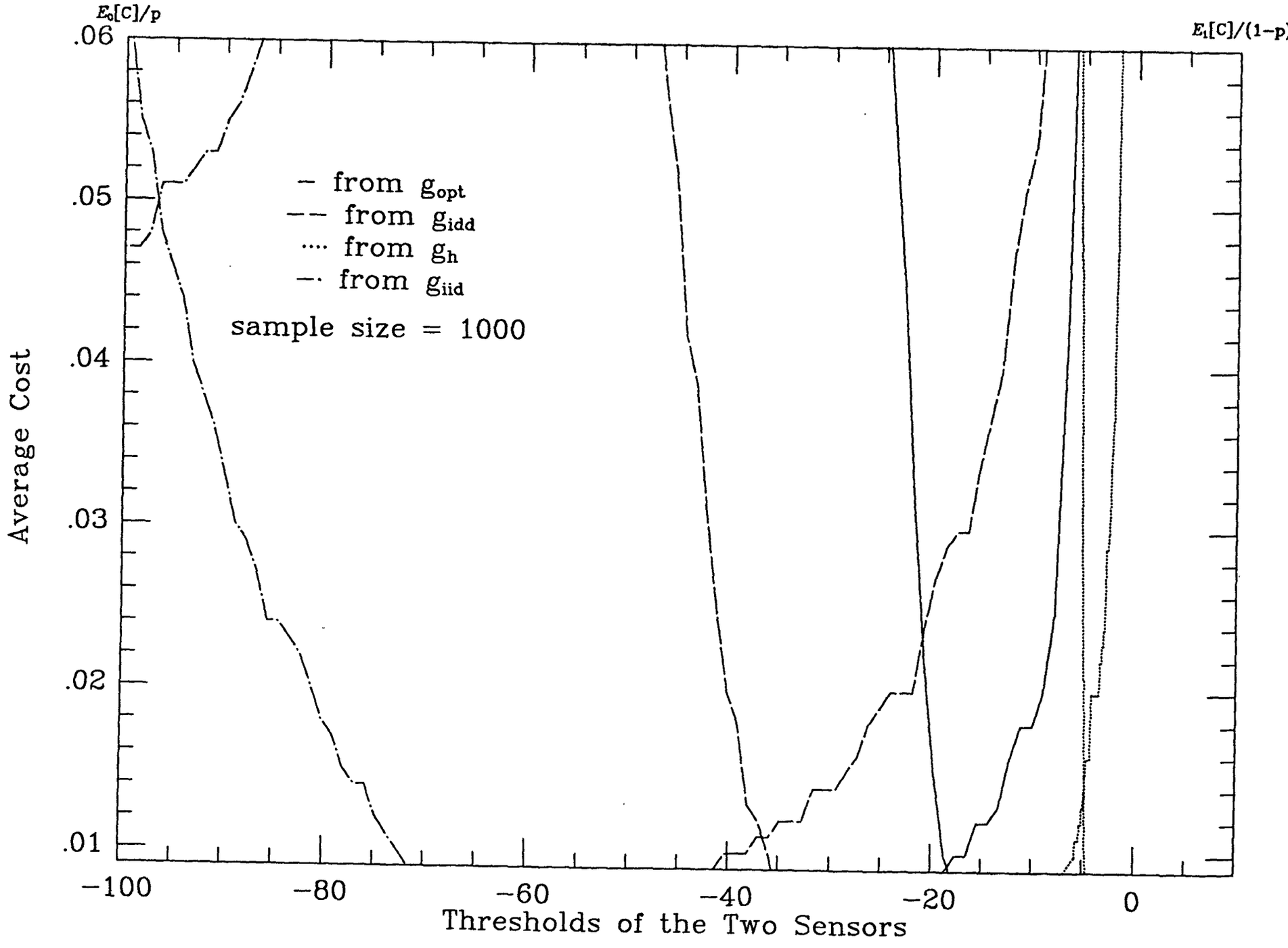

Fig. 4.2 ROCs for Example 4